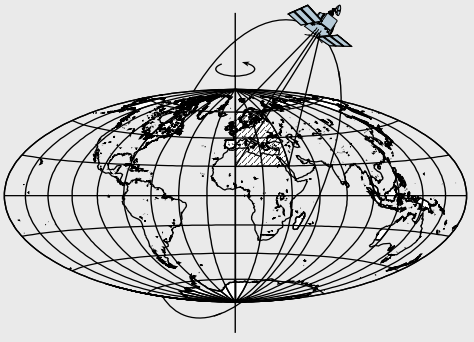
GPS Radio Occultation and the Role of Atmospheric Pressure on Spaceborne Gravity Estimation Over Antarctica

by

Shengjie Ge



Report No. 479

Geodetic Science and Surveying
School of Earth Sciences
The Ohio State University
Columbus, Ohio 43210

July 2006

GPS RADIO OCCULTATION AND THE ROLE OF ATMOSPHERIC PRESSURE ON SPACEBORNE GRAVITY ESTIMATION OVER ANTARCTICA

By

Shengjie Ge

Report No. 479

Geodetic Science and Surveying
School of Earth Sciences
The Ohio State University
Columbus, Ohio 43210

July 2006

ABSTRACT

Dedicated satellite gravity missions are anticipated to significantly improve the current knowledge of the Earth's mean gravity field and its time variable part–climate sensitive gravity signals. They could be measured by the Gravity Recovery and Climate Experiment (GRACE) twin-satellite with sub-centimeter accuracy in terms of column of water movement near the Earth's surface with a spatial resolution of several hundred kilometers or larger, and a temporal resolution of one month or weeks.

To properly recover the time variable gravity signals from space, the gravity measurements require the atmospheric pressure contribution to be accurately modeled and removed. The sparse coverage of measurements makes the weather products less accurate in the southern hemisphere, especially over the Southern Ocean and Antarctica. The asynoptic observation from GPS radio occultation could achieve dense spatial coverage even in remote regions. In this research, we investigate the potential use of GPS occultation to improve the pressure modeling over Antarctica. Atmospheric pressure profiles are retrieved and validated against ECMWF, NCEP and radiosonde observations. Our results show that occultation can provide compatible observations especially in the upper atmosphere. Large standard deviations and biases are found near the ground and in the Antarctic region. GPS occultation in the polar regions is less affected by multipath problem and can penetrate down near the surface. Through an experiment using a 1-D variational (1DVar) approach, we show that the high vertical accuracy of GPS occultation can be propagated down to reduce the uncertainty of surface pressure, indicating that GPS occultation can be expected to have positive impact on the pressure modeling over data-sparse areas after obtaining adequate number of observations (e.g., from Constellation Observing System for Meteorology, Ionosphere & Climate (COSMIC)). We also find that the retrieved profiles could be different due to various assumptions and retrieval algorithms.

Pressure uncertainty degrades the GRACE recovered gravity change. We study the uncertainty of pressure modeling on various temporal scales. Global analysis models show large differences in the Antarctic region. The surface topography may introduce additional biases if it is not well treated. The atmospheric tides are non-negligible and need to be properly considered. The real magnitude of the mismodeled and un-modeled errors in the analysis is hard to evaluate, especially in Antarctica. We simulate the errors sensitive to GRACE using the differences between two global analysis models. Most of the very long wavelength errors are well

captured by GRACE. Their changes in the form of short-period variation increase the errors of the middle to high degree spherical harmonic coefficients. After de-aliasing, middle to high degree coefficients are noticeably improved. The Inverted Barometer (IB) assumption decreases the amplitude of the aliasing error, and the pattern of the RMS difference is slightly changed over land by neglecting the large variations in the Southern Ocean. Our result using more recent ECMWF and NCEP operational analyses shows reduced aliasing effects, which indicates that two models are becoming increasingly close to each other. The model correlation and IB assumption may underestimate the true aliasing error.

The analysis models are validated against the unevenly distributed Automatic Weather Station (AWS) surface pressure observations on the Antarctic continent. Spectral analysis shows that 6-hour analyzed model data can capture most of the power in pressure variations. ECMWF exhibits a much better agreement with AWS than NCEP reanalysis does. Large biases still exist due to the uncertainties of the station elevations. The comparison statistics show strong correlations with the topography with lower standard deviation values in the interior and higher standard deviation values around the coastal area. This result contradicts the distribution derived from the difference between two analysis models, which exhibits large difference in the interior of Antarctica.

We also investigate the influences of different algorithms and assumptions of 2-D or 3-D atmospheric structures on the GRACE atmospheric de-aliasing product. Air density derived from the hydrostatic equation and the equation of state gives slightly different results, and the difference is above the expected GRACE sensitivity. We compare our results with the GRACE atmospheric de-aliasing product and find that the difference is almost below the GRACE sensitivity, although there are differences in the algorithms and we use a relatively low resolution model. We also find that the difference between 3-D hydrostatic formulation and 2-D algorithm is below the GRACE sensitivity. We discover that the atmospheric structure and latitudinal variations of gravity are largely compensated by removing their respective long-term means. Consequently, the 2-D method can greatly reduce the requirements for computational load and data storage. Removing the mean field does not help to reduce the discrepancies between ECMWF and NCEP. If the computational burden is not a concern, using our improved 3-D algorithm can bring a better result. After the full operation of the COSMIC satellites, some major improvement of the pressure modeling over Antarctica is anticipated. A reprocessing of the GRACE data using an improved pressure model could bring us better gravity solutions.

PREFACE

This report was prepared by Shengjie Ge, a graduate research associate in the Geodetic Science and surveying program of the Department of Geological Science at the Ohio State University, under the supervision of Professor C. K. Shum.

This study was partially supported by grants from NASA Interdisciplinary Science Program NAG5-9518, and National Science National Space Weather Program ATM-0418844.

This report was also submitted to the Graduate School of the Ohio State University as a dissertation in partial fulfillment of the requirements for the Ph.D. degree.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION

1.1 Temporal variable gravity and GRACE

The Earth is a dynamic system within which different geophysical processes interact to produce variations with a wide range of temporal and spatial scales. The study of the Earth as a system is a complicated interdisciplinary task. Modeling and separating all these physical processes, including atmospheric and oceanic circulations, redistribution of continental water, cryosphere mass balance, precipitation, snow accumulation and melt, sea level change, luni-solar tides, mantle convection, global isostatic adjustment and plate tectonics, broaden our knowledge of the evolutions of each sub-system and their interactions. Among these physical processes, mass redistribution and transfer are crucial for geophysicists, oceanographers and climate scientists to understand the Earth evolution, climate change, and sea level change. The global climate change signal is directly linked to dynamic processes of the Earth system in the form of mass changes or gravity variations. Integrated mass changes induce perturbations on satellite trajectories. Through accurate tracking of orbital changes, gravity information can be precisely inverted, which indirectly reflects the mass redistribution within the Earth system. Spaceborne dedicated gravity missions have the ability to detect mass exchanges within the interior or outer envelope of the Earth and thus constrain one critical component of the complex Earth system dynamics.

With the satellite gravity missions over the past several decades, the static and the longest wavelength gravity fields have been extensively studied. The temporal gravity change research has been benefiting from the long-term, precise Lageos Satellite Laser Ranging (SLR) observations. The SLR-detected annual variations of the geocenter (degree 1 spherical harmonic coefficients represent the relation between the center of mass and Earth's figure center) are highly correlated with the mass redistribution in the atmosphere, ocean and continental water [*Chen et al.*, 1999]. The Earth's response to the seasonal horizontal mass transportation in the hydrosphere is confirmed from the crustal deformation measured by continuous GPS observations

[*Blewitt et al.*, 2001]. Another term that has been drawing considerable attention is J_2 – the oblateness of the Earth. The secular change of J_2 was first found to be decreasing from SLR data about 20 years ago [*Yoder et al.*, 1983], which was mainly caused by the post glacial rebound. The mantle's viscous relaxation is still going on due to the ice cap melting since the Last Ice Age. A sudden change of J_2 from decreasing to increasing was recently detected by SLR data from multiple satellites [*Cox and Chao*, 2002], and it is asserted to be caused by the surge of sub-polar glacial melting and mass shifts in the Southern, Pacific and Indian Oceans [*Dickey et al.*, 2002]. The Lageo-observed annual and seasonal variations of the low degree coefficients, especially J_2 [*Chao et al.*, 1987; *Chao and Au*, 1991; *Cheng and Tapley*, 1999; *Nerem et al.*, 2000] are mainly induced by the redistribution of the atmospheric [*Trenberth*, 1981; *Trenberth and Guillemot*, 1994; *Trenberth and Smith*, 2005], oceanic and hydrological mass.

The comparisons of the long wavelength signals between the geophysical models and the Lageos observations show very good correlations. However, the results are limited by the low resolution caused by the high altitude of the satellites. Large discrepancies cannot be fully explained using the existing observations. More detailed information is expected for each of these subsystems to identify the missing components, and thus we might have a better understanding of this coupled system. To be more sensitive to the short wavelength signals, the satellite orbit needs to be lowered from around 6,000 km for Lageos to around 250-500 km. The lifetime of such missions, however, is greatly restricted by non-gravitational forces, especially the atmospheric drag.

Among the Low Earth Orbital (LEO) satellites, one is the NASA/DLR Gravity Recovery and Climate Experiment (GRACE) mission [*Tapley et al.*, 2004a,b]. GRACE consists of two identical satellites, orbiting the Earth at an altitude of 500 km and with an inclination of 89°. The separation of the two satellites is approximately 220 km. GRACE was launched in March 2002 in a five year mission to measure climate-sensitive signals in the form of gravity change. The precise orbits of the spacecrafts in the geocentric reference frame are tracked by the onboard Global Positioning System (GPS) receivers in High-Low mode. The two satellites are mutually tracked by the K-band microwave ranging system, which has the advantage of tracking much shorter wavelength signals and is more accurate than the traditional High-Low GPS system. The auxiliary data include the accelerometer data to remove the non-gravitational forces and the star camera data to control the attitude. These subsystems together are able to determine the global gravity field with a spatial resolution from 400 km to 40,000 km approximately every 30 days. GRACE has been anticipating to has the theoretical and practical capabilities to identify the signals from the continental hydrology [*Wahr et al.*, 1998; *Rodell and Famiglietti*, 1999, 2001; *Swenson et al.*, 2003; *Rodell et al.*, 2004a; *Swenson and Milly*, 2006], postglacial rebound and polar ice mass balance [*Wahr et al.*, 2000; *Velicogna and Wahr*, 2002],

deep ocean currents [*Wahr et al.*, 2002], ocean mass redistribution [*Nerem et al.*, 2003], and ocean heat content [*Jayne and Wahr*, 2003]. Early results [*Tapley et al.*, 2004a; *Wahr et al.*, 2004] demonstrated that annual and seasonal variations of water storage can be properly recovered in reasonable agreement with geophysical models for several large drainage basins (e.g., Amazon, and Mississippi) with large signals after applying appropriate smoothing functions. Experiments by *Chen et al.* [2005b] found that 800 km is a suitable radius for the Gaussian smoothing [*Jekeli*, 1981], and could achieve the best RMS value for the difference between GRACE and Global Land Data Assimilation System (GLDAS) [*Rodell et al.*, 2004b] developed at NASA Goddard Space Flight Center. The very low degree spherical harmonics, especially J_2, are not well sensed by GRACE due to its orbital configuration. Replacing these coefficients with those derived independently from Earth rotation and satellite laser ranging improves the agreement for terrestrial water storage [*Chen et al.*, 2005a] and ocean mass variations [*Chambers et al.*, 2004]. The currently released monthly solution error is about 40 times worse than the pre-launch estimate [*Wahr et al.*, 2004]; more improvements are expected after better understanding of the measurement system and various geophysical modelings.

1.2 Atmospheric signals and errors

What GRACE senses is a combined signal from all processes that can perturb the GRACE orbit. After removing non-gravitational effects using the precise accelerometer data, measurements still include the tidal effects (solid Earth, ocean, and pole tides), non-tidal atmospheric, oceanic and hydrological variations, postglacial rebound, ice sheet and glacier mass balance, other un-modeled processes or signals, and noise. To obtain any individual contribution, independent knowledge is required to separate these signals. The candidates of subjects to study largely depend on whether other processes are precisely modeled or measured and their associated error characteristics. Dynamic atmospheric circulation is the most rapidly changing and currently the best measured fluid within the Earth's subsystems. Numerical Weather Prediction (NWP) analysis results can routinely be obtained from the European Center for Medium-range Weather Forecast (ECMWF) and the National Center for Environmental Prediction (NCEP). So the atmospheric pressure change from a NWP model, oceanic mass change from a barotropic ocean model as well as tide models are treated as known, and serve as background models in the data processing. Consequently, the unmodeled signals (e.g. from hydrology) along with the missing and mis-modeled components in the background models are contained in the estimated GRACE solution.

Besides GRACE, atmospheric pressure affects a wide range of other high precision geodetic measurements, such as the tropospheric delay for GPS [*Bevis et al.*, 1992], Inverted Barometer (IB) effects and the tropospheric delay for satellite altimeters [*Fu*

and Pihos, 1994], elastic loading on the Earth surface [*van Dam and Wahr*, 1987; *van Dam et al.*, 1994], and surface gravimetric measurements [*Boy et al.*, 2002]. Accurate atmospheric pressure information is very crucial to improve the quality of these measurements, especially for GRACE. The variation of surface pressure is usually larger than the oceanic and hydrological variations. The SI unit of pressure is Pascal (Pa). One Pascal is the pressure generated by a force of 1 Newton on an area of 1 square meter. In meteorology, the commonly used unit is hectopascal (hPa) or millibar (mbar), where 1 mbar = 1 hPa = 100 Pa = 100 Newton/m^2. Under the IB assumption based on the hydrostatic relation [*Wunsch and Stammer*, 1997], 1 hPa surface pressure error in the seasonal signal translates to 1 cm equivalent water thickness. Because of the nature of the GRACE orbit and the coupling of atmosphere-hydrology-ocean signals, GRACE satellites see a different temporal sampling of the geophysical signals along their orbital path. Any error or uncertainty left in the atmospheric model will inevitably leak into and deteriorate the hydrological or other interesting signals. To recover the hydrological signal with an accuracy of 1 cm equivalent water thickness at a scale of a few hundred kilometers or larger, the atmospheric pressure must be known with an accuracy of 1 hPa or better. *Velicogna et al.* [2001] compared the ECMWF surface pressure with ground measurements in the United States continent and the Arab peninsula over a 30-day period. They found that the error is generally below 0.5 hPa and concluded that the current analyzed surface pressure is sufficient for hydrological purposes. However, the result is not applicable to the Southern Ocean and Antarctica. Over the southern polar region, NWP models suffer from a combination of factors, which include the lack of observations, remote and unreliable communication, steep and not well-known topography, extremely cold weather, and unusual physics in cold environment. Users usually have less confidence on the model outputs in such regions than in the mid-latitudes. The discrepancies between models are larger than in any other area, which has been mentioned or confirmed by many researches [e.g. *Trenberth*, 1992; *van Dam and Wahr*, 1993; *Ponte and Dorandeu*, 2003; *Wahr et al.*, 1998]. Our study [*Ge et al.*, 2002] also shows that the surface pressure uncertainties can easily exceed 6 hPa in Antarctica. A careful examination of surface pressure errors over such region is vital to retrieve the postglacial rebound and ice mass balance signals from GRACE.

The largest uncertainty contributing to the global sea level rise is the stability of the two largest ice sheets: Antarctica and Greenland. An accurate estimation of Antarctic mass balance could effectively reduce the uncertainty of present-day sea level rise. There have been continuous improvements in the understanding of Antarctic mass balance using contemporary remote sensing techniques [*Davis et al.*, 2005]. Because of the location, size and complexity of the Antarctic ice sheet, this region is not adequately measured and each of the remote sensing techniques has its own limitation [*Rignot and Thomas*, 2002]. GRACE, as a supplementary technique, can be used to "weigh" the mass change of the ice sheet. The total ice sheet balance

includes the accumulation (surface mass balance), sub-ice-shelf melt, and iceberg calving. The accumulation generally includes the precipitation, evaporation, blowing snow, and run-off. The spatial variability of Antarctic precipitation is very large and generally follows the topography. The precipitation in the interior is less than 50 mm/year (equivalent water thickness) due to lack of forcing and topographic blocking, but in some coastal regions it could exceed 500 mm/year [*van Lipzig et al.*, 2002]. Blowing snow is relatively small in magnitude and contributes generally less than 5% across the boundary of the continent. The accuracy of the precipitation from models is worse than 5%, and therefore, blowing snow is negligible. The run-off can be neglected as well, since most of the melt water in Antarctica refreezes close to where it is formed [*Liston and Winther*, 2005]. Most of the ice lost in Antarctica is from basal (sub-ice-shelf) melt and iceberg calving, which were shown to be substantial and hard to measure [*Rignot and Thomas*, 2002]. The average snow accumulation is about 160 mm/year over the whole Antarctica [*Vaughan et al.*, 1999], and distributes unevenly both in space and time. The large uncertainty in the pressure could make the secular, annual or seasonal ice/snow variations difficult or even impossible to determine. The spatial distribution of determined ice/snow variations is also affected by the distribution and variation of the pressure error.

Surface pressure is not the exact representation of the total atmospheric mass. The vertical structure of the atmosphere may affect the computation of the atmospheric mass variation. *Swenson and Wahr* [2002] investigated the effect of the thickness of the atmosphere-to-mass signal using a simplified two-layer model and drew the conclusion that surface pressure could achieve comparable accuracy with respect to the multi-layered model and significantly reduce the computational requirement. However, *Boy and Chao* [2005] suggested that the difference between 3-D and 2-D computations was non-negligible up to spherical harmonic degree of 15-20 in the presence of the expected GRACE sensitivity. This conclusion might also be sensitive to the data used. It is required to further validate the influences of these two statements on the model-pressure calculated mass variations.

In addition, temporal aliasing effect, associated with the non-tidal, short-period atmospheric and oceanic variations, is another error source for the gravity mapping [*Gruber et al.*, 2001] from GRACE. Neglecting these errors could cause substantial errors [*Han et al.*, 2004; *Thompson et al.*, 2004]. It is usually difficult to determine the real aliasing error contained in the solution. Simulation results largely depend on the different variations between any two analyzed models (e.g., between ECMWF and NCEP). The IB assumption may also affect the magnitude and the distribution of the aliasing error.

The best way to obtain accurate pressure data is to extend the current meteorological network. This suggestion is sometimes impractical and expensive. To set up and maintain a US-like barometric network in the Southern Ocean and Antarctica is

almost impossible. GPS radio occultation has the potential to improve the pressure field in those remote, hard-to-access areas.

1.3 GPS radio occultation

Radio occultation, as a new remote sensing technique to explore the atmosphere, was first used in NASA's planetary missions to probe the planetary atmosphere [*Fjeldbo and Eshleman*, 1969]. This technique has been applied for the limb-sounding of the Earth's atmosphere from Low Earth Orbiters (LEOs, 400-1300 km altitude). The GPS receivers carried by the LEO acquire signals from the constellation of GPS satellites at 20,000 km altitude. The GPS radio occultation concept has been successfully demonstrated for the first time by the GPS/MET (MicroLab-1) experiment in 1995 [*Ware et al.*, 1996]. The fundamental principle is that transmitted GPS signals are delayed when LEO is setting into or rising above the Earth's atmosphere. This delay could be measured in the form of radar Doppler shift and a bending angle of the radar path. Finally, the atmospheric refractivity, temperature, pressure, air density as well as water vapor can be retrieved by assuming knowledge of appropriate boundary conditions.

After the success of the GPS/MET [*Ware et al.*, 1996] proof-of-concept mission, the current German CHAllenging Minisatellite Payload (CHAMP, launched in July 2000) [*Reigber et al.*, 2002] and Argentina's SAC-C (launched in November 2000) carrying a new generation of GPS flight receiver ("Blackjack") provide quasi-continuous GPS occultation measurements. In addition, the U.S.-German GRACE mission (launched in March 2002), and the Taiwan-U.S. multi-satellite Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC, launched in April 2006) [*Anthes et al.*, 2000] mission will provide unprecedented opportunities to continuously observe the Earth's atmosphere by the GPS radio occultation technique. The lower troposphere is one of the main foci in atmospheric science. Understanding the Planetary Boundary Layer (PBL), water vapor circulation, heat and energy transportations, etc., is critical to the further improvement of weather forecasting and climate research. Because of the rapid change of the moisture content, lower troposphere inhomogeneities, and the multipath problem, difficulties with the signal penetration through the lower troposphere are still a big challenge in both the robustness of the retrieval algorithm and the hardware technology.

The GPS radio occultation technique has been described in numerous literatures [e.g. *Gorbunov and Sokolovsky*, 1993; *Melbourne et al.*, 1994; *Kursinski et al.*, 1996; *Gorbunov et al.*, 1996]. It can provide profiles with a high vertical resolution [*Kursinski et al.*, 1996] and has the best result at 5-30 km altitude. In this range, the derived temperature agrees with ECMWF better than 1.5-2°K, and the accuracy continues to be improved from GPS/MET [*Rocken et al.*, 1997], CHAMP [*Wickert et al.*, 2001] and GRACE [*Beyerle et al.*, 2005]. Geopotential heights derived from

early GPS/MET data agrees with that from ECMWF within 20 gpm (geopotential meter, see Appendix A) [*Leroy*, 1997]. The GPS occultation measurements are more suitable for climate study because of their comparable accuracy with radiosonde measurements and no need for heterogeneous instrument calibration [*Kuo et al.*, 2005]. The retrieval methods are further refined to improve the accuracy in the lower troposphere, which include using amplitude data [*Sokolovskiy*, 2000], radioholographic method [*Hocke et al.*, 1999; *Gorbunov et al.*, 2000], back-propagation method [*Gorbunov*, 2001], canonical transformation method [*Gorbunov*, 2002], full spectrum inversion [*Jensen et al.*, 2003] and other forms of variations. Like other meteorological observations, GPS radio occultation measurements can be assimilated into the NWP system to generate an "analyzed" field to correct and initialize the forecast system [*Eyre*, 1994]. Active researches and experiments have been carried out for the understanding of the impact of this new data type to the weather forecast system using the limited number of observations from GPS/MET [*Zou et al.*, 1999; *Liu et al.*, 2001] and CHAMP [*Poli et al.*, 2003; *Zou et al.*, 2004; *Healy et al.*, 2005]. Although some studies show that the impact was not significant [e.g. *Poli et al.*, 2003], there are positive improvements on the forecast errors, especially in the southern hemisphere [*Healy et al.*, 2005].

By using the fully deployed GPS and LEO constellations, radio occultation provides an alternative way to obtain atmospheric data for all weather conditions and nearly uniform global coverage, especially for data-sparse regions, e.g., the Southern Ocean and Antarctica. The newly launched 6-satellite COSMIC constellation is expected to have 2,500 soundings in 24 hours. These measurements represent a valuable data source for climate and meteorological researches and for better understanding of the Earth's atmosphere. In addition to applications in atmosphere science and meteorology, GPS occultation data can also be used to improve the global surface pressure field to exploit the accuracy of dedicated gravity missions for the recovery of mean and time variable gravity field.

1.4 Outline of the study

In this research, we mainly focus on two topics. One is the implementation of the methods for accurate GPS-occultation profile retrievals and the discussion of their current limitations. We will describe in detail the algorithms used to develop our processing software. The main difficulty for occultation retrieval is related to the multipath problems at the lower troposphere, which is primarily caused by the rapid variations of the moisture content in the lower atmosphere. Radio holographic methods are introduced to handle the multipath problem. We pay special attention to the properties of occultation profiles over Antarctica, since the polar continental air is relatively dry, and the profile retrieval is relatively simpler than in the tropical region where there is greater interference from the water vapor. However, the physics of the

cold-region atmosphere is distinct from that of middle to low latitude regions, and warrants special attention. We intend to assess the potential impact of using the radio occultation measurements from LEOs (CHAMP and SAC-C) to improve the accuracy of the atmospheric loading for the GRACE temporal gravity field estimation. The retrieved profiles can also be used as independent measurements to evaluate NWP models since most of the GPS-occultation profiles have not been assimilated into the NWP. For the Automatic Weather Station (AWS) observations, some of these ground data have already been used in the analysis models, which may make comparison results optimistic. By examining the statistical information in the comparisons of radio occultation with analysis models as well as inter-occultation and inter-model comparisons, we try to reveal the benefit and weakness of each data type for better understanding and utilizing each data type. With the launches of new missions such as COSMIC, we would expect a better performance of NWP models over data-sparse regions after including these new occultation data.

In the second part of the research, we investigate the role of atmospheric pressure, its structure, variation and underlying mechanism responsible for the contribution of the atmospheric errors to the time-dependent spaceborne gravity estimation. At present, users have the least confidence in the analyzed products over the southern polar region. Selecting the best dataset from different analyzed models needs careful evaluation of their respective performances. In this study, we evaluate different meteorological analysis products (e.g., ECMWF and NCEP, operational and re-analysis) over the Antarctic region using surface AWS observations. The statistics can be used as a valuable indicator for the model performance, though this data type may not be completely independent. The discrepancy between analysis models and their impact on time variable gravity field will be discussed. The mathematical modeling is another important aspect that we focus on. Since there are many assumptions and approximations in the process of converting pressure to mass, we need to evaluate their respective influences on the gravity estimation. In addition, we investigate the effects of different temporal and spatial resolution data, the interpolation method, the topography, IB, and 2-D/3-D modeling on the monthly gravity solution. The goal is to assess the potential of improved atmospheric model for more accurate gravity field solution using GRACE data over the Antarctic region.

1.5 Organization of the dissertation

The rest of the dissertation is organized as follows.

Chapter 2 introduces the basic principle of the GPS occultation. The method using geometric optics is discussed in detail. The resolution of this method and processing procedures are summarized as well.

Chapter 3 emphasizes the algorithms of the GPS occultation in handling the multipath problem in the lowest region of the troposphere. The limitation of the

geometric optics method is addressed and various radio holographic methods (using wave equations) are discussed and compared.

Chapter 4 gives the GPS occultation retrieval results. Its validation against the analyzed models, radiosonde measurements and inter-occultation comparisons are investigated. The special characteristics of the GPS occultation in Antarctica is addressed as well.

Chapter 5 discusses the uncertainties and other effects on the atmospheric pressure model. The effects of the topography, atmospheric tide, inverted barometer, and temporal aliasing on the time variable gravity estimations based on different analyzed models are examined.

Chapter 6 evaluates the atmospheric error using the AWS observations over Antarctica, and the implication on model accuracy is also discussed in this chapter.

Chapter 7 systematically introduces the mathematical model for atmospheric dealiasing. The algorithms based on the hydrostatic equation and the equation of state, 2-D/3-D atmospheric structure, as well as the influence of the mean field and the use of different NWP models are investigated.

Chapter 8 summaries the results and the conclusions in this study, along with discussions and outlook.

CHAPTER 2

GPS RADIO OCCULTATION PROFILE RETRIEVAL USING GEOMETRIC OPTICS

2.1 Introduction

In geometric optics [*Born and Wolf*, 1993], the propagation path of electromagnetic (EM) waves can be assumed as dimensionless rays connecting the transmitter and the receiver. The solutions are based on the assumption of a spherically symmetric distributed atmosphere described by Snell's law. The change of the Doppler frequency along the orbit of the Low Earth Orbiter (LEO) satellite is used to derive the incident ray direction (or phase front slope) at each point on the orbit. The bending angle of the ray is geometrically obtained thereafter. This method is only valid when no more than one ray arrives at every observation point on the LEO trajectory. The vertical profile of atmospheric refractivity is related to the bending angle through the Abel Transformation [*Fjeldbo et al.*, 1971]. Afterward, the atmospheric pressure, temperature, and humidity can be inverted from the refractivity through well established formulas. This method is relatively simple and straight forward. Usually, we apply this method in regions with no multipath effect such as in the upper troposphere and stratosphere.

2.2 The radio signal refraction in the atmosphere

The propagation speed of the electromagnetic signal in vacuum is the speed of light. When the same signal passes through the Earth's atmosphere, it interacts with the neutral or charged atoms, molecules, or other particles, and the direction and speed of the wave propagation are changed. The ratio of the speed of the wave propagation in vacuum (c) to the speed of propagation in the medium (v) is called the refractive index (n) of the medium:

$$n = \frac{c}{v} \tag{2.1}$$

In the neutral atmosphere and for frequencies less than 30 GHz, the phase velocity and wave number are not dependent on the frequency of the wave. This is known as a non-dispersive medium. This region includes the troposphere and stratosphere, ranging from 0 to 50 km altitude. The refractivity in this region is mainly due to the dipole-type dry air and water vapor molecules, which are polarized by the incident electric field, such as ultraviolet electronic transitions.

2.2.1 The ray trajectory and refractivity

In geometric optics, the propagation of the electromagnetic waves is treated in terms of rays. The trajectory of the ray connecting two points when passing through a region with varying refractive index obeys Fermat's principle of least time:

$$\int_1^2 n(\mathbf{r})ds = min \tag{2.2}$$

It is usually impossible to retrieve the three-dimensional distribution of the refractivity in the vicinity of the ray perigee (defined in Figure 2.3) from the one dimensional measurements along the trajectory of the receiving satellite. The problem can be simplified by considering the fact that the characteristic horizontal scale of meteorological fields is much bigger than the horizontal displacement of the ray perigee that crosses the atmospheric layers [*Gorbunov and Sokolovsky*, 1993]. This implies that the bending of the ray is generated dominantly by the vertical gradient of the refractive index in the occultation plane (the plane is formed by connecting two satellite points and the refraction center). The gradient of the refractive index normal to the occultation plane has negligible effect on the Doppler measurements. Thus, we assume the ray path resides only in the vertical occultation plane. The gradient of n is in the radial direction by assuming that the local refractivity distribution is spherically symmetric in the vicinity of the ray perigee, such that

$$n(\mathbf{r}) = n(r) \tag{2.3}$$

Substitute the above equation into equation (2.2), the ray path is described as Snell's law:

$$rn(r)\sin(\phi) = p \tag{2.4}$$

where ϕ is the angle between the ray path and the refractivity gradient vector. p is a constant for a given ray trajectory. It is usually called *impact factor* or *impact parameter*. If the ray passes through a spherical symmetric refractive field and lies within the occultation plane, the change of the ray direction can be described by a polar coordinate system in such a plane with the origin at the center of the refractive field. The refractive index is only a function of the radius r, since the atmospheric density is distributed symmetrically in layers (like an onion) and its value increases as r decreases. The geometry of the total bending of a ray is shown in Figure 2.1. ϕ is

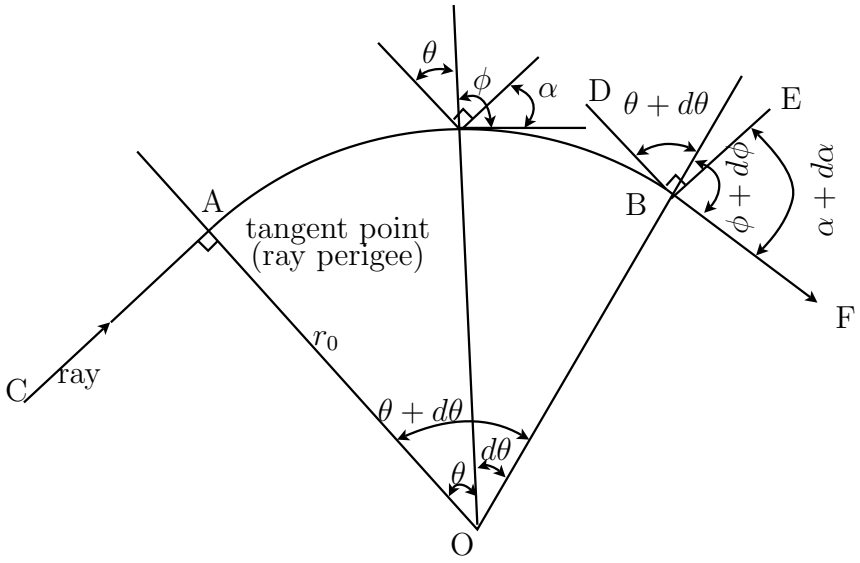


Figure 2.1: The geometry of the refraction of the ray.

the angle between the ray path and the local radial direction, which can be expressed using the polar coordinates

$$\tan(\phi) = \frac{r d\theta}{dr} \tag{2.5}$$

From Figure 2.1, the total bending angle α of the ray is

$$\alpha = \theta + \phi - \frac{\pi}{2} \tag{2.6}$$

Differentiating equation (2.6) and combining with equation (2.5), we obtain

$$d\alpha = d\theta + d\phi \tag{2.7}$$

The expressions for $d\theta$ and $d\phi$ are needed to compute the change of the bending angle $d\alpha$.
From Snell's law, $\tan\phi$ can be derived in another form from equation (2.4) as:

$$\tan(\phi) = \frac{p}{\sqrt{(nr)^2 - p^2}} \tag{2.8}$$

Combining the above two equations (2.5) and(2.8), $d\theta$ is written as:

$$d\theta = \frac{p dr}{r\sqrt{(nr)^2 - p^2}} \tag{2.9}$$

For the expression of $d\phi$, we can differentiate equation (2.4),

$$r\sin\phi dn + n\sin\phi dr + nr\cos\phi d\phi = 0 \tag{2.10}$$

Rewrite above as:

$$d\phi = -\frac{\tan\phi(rdn + ndr)}{nr} \tag{2.11}$$

Substituting equation (2.8) into the above equation for $\tan\phi$, yields,

$$d\phi = -\frac{p(rdn + ndr)}{nr\sqrt{(nr)^2 - p^2}} \tag{2.12}$$

From equations (2.9) and (2.12), $d\alpha$ in equation (2.7) can be written as:

$$d\alpha = \frac{pdr}{r\sqrt{(nr)^2 - p^2}} - \frac{p(rdn + ndr)}{nr\sqrt{(nr)^2 - p^2}} = -\frac{p}{\sqrt{(nr)^2 - p^2}}\frac{d\ln(n)}{dr}dr \tag{2.13}$$

Integrating equation (2.13) along the entire ray path, the total bending angle is obtained as:

$$\alpha = \int_0^{\alpha'} d\alpha = -2p\int_{r_0}^{\infty} \frac{1}{\sqrt{(nr)^2 - p^2}}\frac{d\ln(n)}{dr}dr \tag{2.14}$$

The "2" in the above equation is introduced due to the assumption of the spherical symmetry of the atmosphere, i.e., the ray is symmetric about the tangent point. Equation (2.14) gives us a forward relationship to compute the bending angle profiles when we have $n(r)$ along the ray path. It reflects that the bending angle is caused by the radial variations of the refractivity. This equation is usually called the *Abelian integral equation*.

2.2.2 The Abelian inversion

The Abelian integral equation (2.14) can be inverted to express the refractive index as a function of the bending angle and the impact factor. To facilitate the derivation of the formula, an auxiliary variable is defined as:

$$x = nr \tag{2.15}$$

Equation (2.14) can be rewritten as:

$$\alpha(p) = -2p\int_{x=p}^{x=\infty} \frac{1}{\sqrt{x^2 - p^2}}\frac{d\ln(n)}{dx}dx \tag{2.16}$$

Dividing both sides of equation (2.16) by the factor $\sqrt{p^2 - p_0^2}$ and integrating with respect to p from p_0 to ∞, we obtained the following relation. p_0 is the impact factor

corresponding to the radius r_0 at the tangent point:

$$\int_{p=p_0}^{p=\infty} \frac{\alpha(p)}{\sqrt{p^2-p_0^2}}dp = -\int_{p=p_0}^{p=\infty} \frac{2p}{\sqrt{p^2-p_0^2}}\left[\int_{x=p}^{x=\infty} \frac{1}{\sqrt{x^2-p^2}}\frac{d\ln(n)}{dx}dx\right]dp \tag{2.17}$$

Changing the order of the integration and choosing the proper limits corresponding to the same integral region on the right side of the equation, yields:

$$\int_{p=p_0}^{p=\infty} \frac{\alpha(p)}{\sqrt{p^2-p_0^2}}dp = -\int_{x=p_0}^{x=\infty} \frac{d\ln(n)}{dx}\left[\int_{p=p_0}^{p=x} \frac{2p}{\sqrt{(p^2-p_0^2)(x^2-p^2)}}dp\right]dx \tag{2.18}$$

Let us introduce another auxiliary variable u as a function of p,

$$u(p) = \frac{p^2-p_0^2}{x^2-p_0^2} \qquad \text{and} \qquad du = \frac{2pdp}{x^2-p_0^2} \tag{2.19}$$

By substituting (2.19) into the inner integral expression in equation (2.18), it becomes:

$$\int_{p=p_0}^{p=x} \frac{2p}{\sqrt{(p^2-p_0^2)(x^2-p^2)}}dp = \int_{u=0}^{u=1} \frac{du}{\sqrt{u-u^2}} = \arcsin(2u-1)|_{u=0}^{u=1} = \pi \tag{2.20}$$

Then, equation (2.18) is simplified as:

$$\int_{p=p_0}^{p=\infty} \frac{\alpha(p)}{\sqrt{p^2-p_0^2}}dp = -\pi\int_{x=p_0}^{x=\infty} \frac{d\ln(n)}{dx}dx = \pi\ln n(r_0) \tag{2.21}$$

Rearranging equation (2.21), yields:

$$n(r_0) = \exp\left[\frac{1}{\pi}\int_{p=p_0}^{p=\infty} \frac{\alpha(p)}{\sqrt{p^2-p_0^2}}dp\right] \tag{2.22}$$

This is the classical *inverse Abel transformation*. The refractivity can be derived from the bending angles and the impact factors in the occultation plane.

2.3 The relationship between refractivity and atmopheric properties

The gas contents and properties contribute to the refractive index n in the neutral atmosphere. Since the value of the refractive index is very close to 1, it is usually more convenient to use the refractivity N, which is defined as:

$$N = (n-1)\times 10^6 \tag{2.23}$$

Generally, different gas species contribute differently to the refractivity for radio frequencies. In a neutral atmosphere, the refractivity of the radio waves that pass through the atmosphere is expressed as a function of the pressure, the temperature, and the humidity content. A common form of the formula is:

$$N = k_1 \frac{P_d}{T} Z_d^{-1} + \left(k_2 \frac{P_w}{T} + k_3 \frac{P_w}{T^2} \right) Z_w^{-1} \tag{2.24}$$

where Z_d and Z_w are the compressibility factors for the dry and moist air respectively. They are used to account for the departure of the air from the ideal gas, and their magnitudes are usually less than 0.001 for the typical atmospheric condition. P_d is the dry atmospheric pressure in hPa, T is the atmospheric temperature in kelvin, and P_w is the partial pressure of water vapor. k_1, k_2, and k_3 are empirically determined coefficients.

In meteorology, the equation for refractivity is often given as:

$$N = c_1 \frac{P}{T} + c_2 \frac{P_w}{T^2} \tag{2.25}$$

where

$$c_1 = k_1 \tag{2.26}$$

$$c_2 = (k_2 - k_1)T + k_3 \tag{2.27}$$

Equation (2.25) is also known as the Smith-Weintraub equation [*Smith and Weintraub*, 1953]. Where P is the atmospheric pressure in hPa. The values for the constant coefficients are: $c_1 = 77.60, c_2 = 3.73 \times 10^5$. These constant values are usually valid for radio frequencies less than 20 GHz.

2.4 Derivation of the atmospheric properties

In the neutral atmosphere, both the dry and moist air contribute to the total refractivity. In the lower part of the atmosphere, especially in tropical regions, the rapidly changing moisture content is a big contributor to the refractivity, and needs special considerations. For the regions where there's no moist air or the moist air has a negligible effect on the refractivity, the atmospheric parameters can be easily derived using a combination of equation (2.25) and the ideal gas law (or the Equation of State):

$$P = \frac{\rho R T}{m} = \rho R_m T \qquad \text{with} \qquad R_m = \frac{R}{m} \tag{2.28}$$

Here ρ is the air density, R is the universal gas constant ($R = 8.3144\ Jmol^{-1}k^{-1}$), m is the mean molecular mass of the gas, and R_m is the specific gas constant, which depends on the molecular weight of the air.

Air is composed of dry and moist air. We can rewrite the ideal gas equation as:

$$P = P_d + P_w = \frac{\rho_d RT}{m_d} + \frac{\rho_w RT}{m_w} = \frac{RT}{V}\left(\frac{M_d}{m_d} + \frac{M_w}{m_w}\right) \tag{2.29}$$

Where M_d is the mass of the dry air, M_w is the mass of the moist air, m_d and m_w are the molecular weight of the dry and moist air respectively. It is also obvious that:

$$\rho = \frac{M}{V} = \frac{M_d + M_w}{V} \tag{2.30}$$

Eliminating V in equation (2.29) by combining with equation (2.30):

$$P = \rho RT\left(\frac{M_d}{m_d} + \frac{M_w}{m_w}\right)\frac{1}{M_d + M_w} = \rho\frac{R}{m_d}T\frac{1 + q/\epsilon}{1 + q} = \rho R_d T\frac{1 + q/\epsilon}{1 + q} \tag{2.31}$$

In the above equations, we define $q = M_w/M_d$, $\epsilon = m_w/m_d$ and $R_d = R/m_d$. A good approximation for equation (2.31) is given as:

$$P = \rho R_d T(1 + 0.608q) = \rho R_d T_v \qquad \text{with} \qquad T_v = T(1 + 0.608q) \tag{2.32}$$

This is another form of the *Equation of State*. T_v in the equation is called the virtual temperature. With this virtual correction, the presence of water vapor is compensated. Therefore, we can use the dry air constant in the Equation of State.

2.4.1 Dry air density

If only the dry part of the air is considered, equation (2.25) becomes:

$$N = c_1\frac{P}{T} \tag{2.33}$$

Combining with equation (2.28), the dry air density profile is given as:

$$\rho_d(z) = \frac{m_d}{Rc_1}N(z) \tag{2.34}$$

Where z is the height above a reference surface (ellipsoidal height or mean sea level height).

2.4.2 Dry air pressure

Usually, for a large volume of air, in the vertical direction, the gravity and the pressure gradient force are in balance. It is called the hydrostatic equilibrium. The expression is:

$$\frac{\partial P}{\partial z} = -\rho(z)g(z) \tag{2.35}$$

$g(z)$ is the gravity acceleration at height z. Substituting equation (2.34) into the above equation and integrating it, the pressure is represented as a function of height:

$$P(z) = \int_z^\infty g(z')\rho_d(z')dz' = \frac{m_d}{Rc_1}\int_z^\infty g(z')N(z')dz' \tag{2.36}$$

2.4.3 Dry air temperature

Using either equation (2.33) or equation (2.28) again, the temperature profile is given as:

$$T(z) = \frac{m_d}{RN(z)}\int_z^\infty g(z')N(z')dz' \tag{2.37}$$

We can directly derive the temperature profile from the refractivity profile under the assumption of the non-existence of the moist air.

2.4.4 Air Humidity

In the GPS occultation, the retrieval of the pressure and temperature from the refractivity also requires the knowledge of the water vapor pressure P_w. In the upper troposphere and stratosphere, the atmosphere can be assumed to be dry with negligible error. However, in the lower troposphere and the warmer tropical regions, where the contribution of the wet component to the refractivity is significant, it is sometimes up to 30% of the total refractivity. The water vapor cannot be easily separated from the dry component. Neglecting the moist component will cause a big error. The extra moisture term in the equations makes the recovery of the air temperature, density, and pressure not unique. This ambiguity would not allow us to distinguish between the contributions of the dry and moist terms in the total refractivity unless independent or *a priori* information is provided. Another important reason is that the hydrostatic equation only works with the total pressure. In order to compute the water vapor pressure, an accurate and independent estimate of temperature must be known. This may come from the NCEP or ECMWF meteorological analyses or forecast models. With the prior knowledge, the water vapor can be computed from an iterative process [e.g. *Gorbunov et al.*, 1996]. This algorithm suffers from a high sensitivity to even small errors in the analyzed temperatures, resulting in large uncertainties of the derived water vapor profiles [*Marquardt et al.*, 2001]. A more elaborate method to retrieve both the water vapor and temperature in an optimal way based on the variational principle will be introduced in a later chapter.

By substituting ρ in equation (2.32) to equation (2.35), it becomes:

$$\frac{dP}{P} = -\frac{g}{R_d T_v} \tag{2.38}$$

Integrating the above equation from z to z^*, yields:

$$P(z) = P(z^*) \exp\left(\int_z^{z^*} \frac{g(z')}{R_d T_v(z')} dz' \right) \tag{2.39}$$

If $P(z)$ is known, $P_w(z)$ can be derived from equation (2.25) as:

$$P_w(z) = T^2(z) \frac{N(z) - c_1 \frac{P(z)}{T(z)}}{c_2} \tag{2.40}$$

The specific humidity is defined as:

$$q(z) = \frac{\epsilon P_w(z)}{P(z) - (1-\epsilon) P_w(z)} \tag{2.41}$$

where ϵ is usually given the value of 0.622.

Summarizing, for a given temperature profile, the iterative procedure of calculating the specific humidity is shown in Figure 2.2.

The mean spatial distribution of the water vapor is similar to the spatial distribution of the temperature, decreasing from the equator to the pole. This would make the pressure retrieval in the polar region more easy, accurate and less affected by the water vapor. Some researchers suggest that water vapor does not need to be considered until the temperature is larger than 250°K [*Kursinski et al.*, 1996]. Therefore, in most of the cases, it is fair to neglect the water vapor when retrieving the profiles over the cold polar regions.

2.5 The excessive phase measurement of the GPS occultation

The Global Positioning System (GPS) consists of a constellation of 24 satellites each with an altitude of 20,200 km. These satellites are evenly distributed in six orbital planes inclined 55 degrees each. The orbital period is around 11 hours 58 minutes. The GPS carrier phase signals are transmitted on two L band frequencies: L1–1.57542 GHz and L2–1.2276 GHz. A LEO carrying a GPS receiver can receive signals from the GPS constellation. Normally, we can express the complex signal received by the circuit of the GPS receivers as $u(t) = A(t) \exp(i\Phi(t)) = A(t) \exp(ik\Psi(t))$. $A(t)$ is the amplitude, $k = 2\pi/\lambda$ is the wavenumber, and $\Psi(t)$ is the length of the ray path (unit: meter) from the transmitter to the receiver, it is also called the eikonal of the ray. $\Phi(t) = k\Psi(t)$ is the phase of the ray (unit: radian). In the subsequent derivations, we sometimes treat Ψ as the phase delay expressed in the unit of length. When a GPS signal is penetrating the neutral atmosphere, it is bent dominantly by the vertical gradient of the refractive index field. The excessive phase delay $\Delta\Psi_{1,2}$

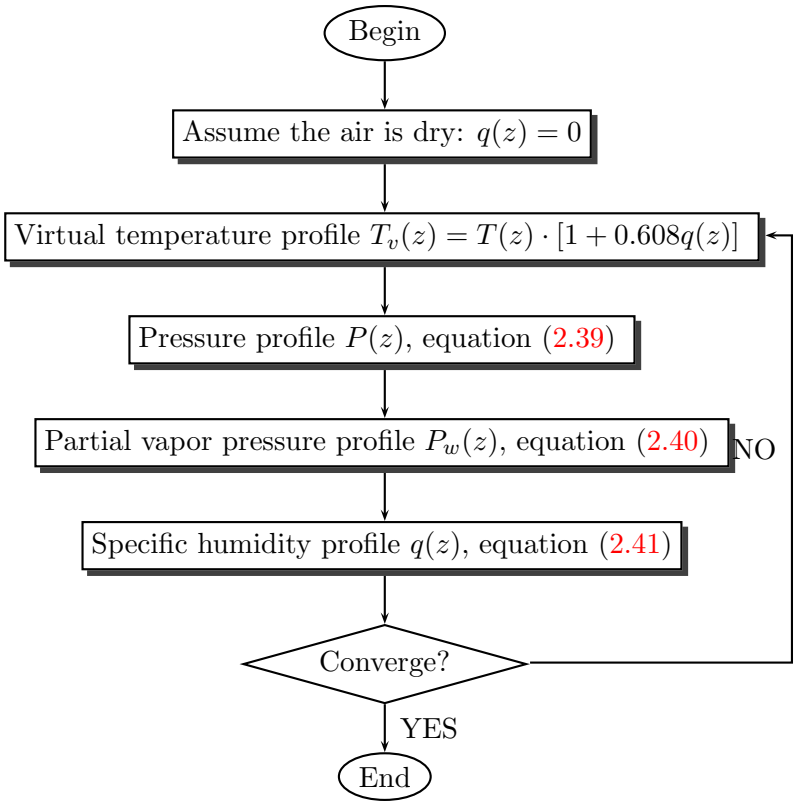


Figure 2.2: The iterative process of computing a humidity profile.

(unit: meter) from the position of the transmitter a to the position of the receiver b is defined as:

$$\Delta\Psi_{1,2} = (\Delta\Psi_{1,2})_{geometric} + (\Delta\Psi_{1,2})_{optical} = \int_a^b n(\Psi_{1,2}) d\Psi_{1,2} - \Psi_0 \tag{2.42}$$

where $n(\Psi_{1,2})$ is the group refractivity index along the ray path for L1 and L2 frequencies, Ψ_0 is the length of the straight line connecting the GPS satellite and LEO. The total atmospheric excessive phase delay $\Delta\Psi_1$ and $\Delta\Psi_2$ in the occultation link can be separated to two effects:

Geometrical effect The increased length of the ray path due to the bending (curvature) of the ray in vacuum.

$$(\Delta\Psi_{1,2})_{geometric} = \int_a^b d\Psi_{1,2} - \Psi_0 \tag{2.43}$$

Optical effect the increase of the optical ray path due to traveling through a medium with a refractive index larger than 1 in both the neutral atmosphere and the ionosphere.

$$(\Delta\Psi_{1,2})_{optical} = \int_a^b (n(\Psi_{1,2}) - 1) d\Psi_{1,2} \tag{2.44}$$

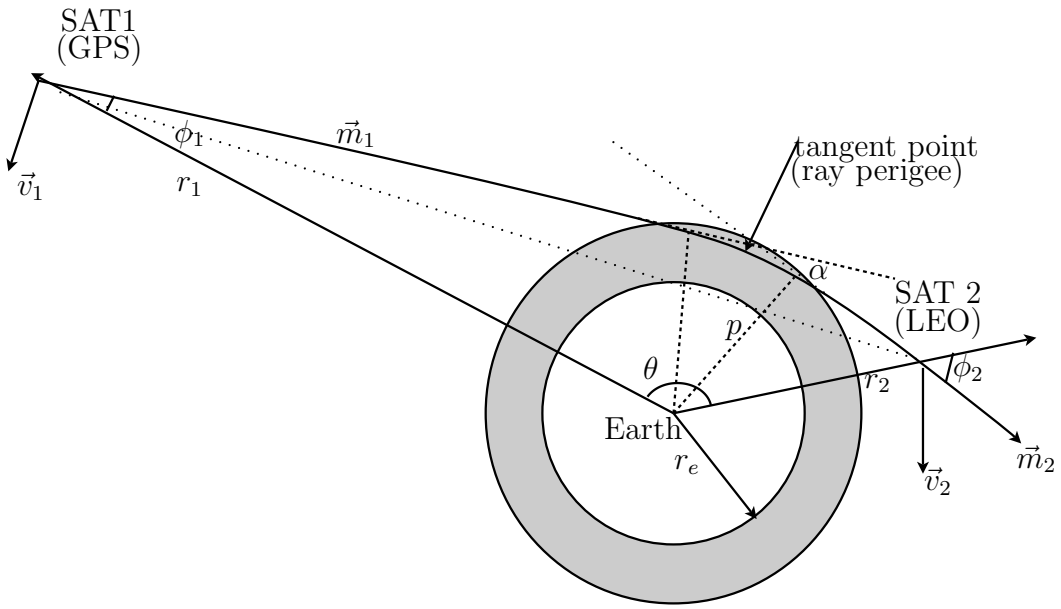


Figure 2.3: The geometric map of GPS occultation.

As we already know from the previous section, through the Abelian inversion, the refractivity profile can be derived from the bending (refraction) angle α and the impact factor p of the ray trajectory (see Figure 2.3). We usually call the bending angle and the impact factor the fundamental measurements of the GPS occultation. The point on the ray path, which is the closest to the Earth's surface, is called the ray perigee or tangent point. The impact factor p is defined as the distance between the ray perigee and the local center of curvature of the ray path.

To derive the bending angle and impact factor, the observed atmospheric excessive phase will first be needed. A double difference technique is used to eliminate satellite clock (including the clock error of the GPS transmitter and the GPS receiver onboard the LEO) errors and to derive the atmospheric excess phase of the occultation link. The double difference technique can also be used to remove the Selective Availability

(SA) error, which was added to the GPS satellites' clock and ephemeris to degrade the positioning accuracy for civilian users, for the historic GPS/MET data. SA was officially turned off on May 1, 2000. The carrier phases (L1 and L2) from the occulting GPS satellite is sampled in 50 Hz by the GPS receiver on the LEO, along with the measurement from a second, referencing GPS satellite, are synchronized with interpolated 1 Hz data from a global distributed fiducial network of GPS ground receivers.

Alternatively, as a result of the termination of the Selective Availability (SA) mode of the GPS (resulting in enhanced GPS clock stability), the application of single differencing technique for precise occultation processing becomes feasible and was demonstrated by *Wickert et al.* [2002]. Additional data needed for the excess phase derivation are the precise orbit information (position and velocity) of the GPS and LEO satellites. Current orbit determination can achieve centimeter accuracy in position and 0.1 mm/s in velocity, which is adequate for the GPS occultation.

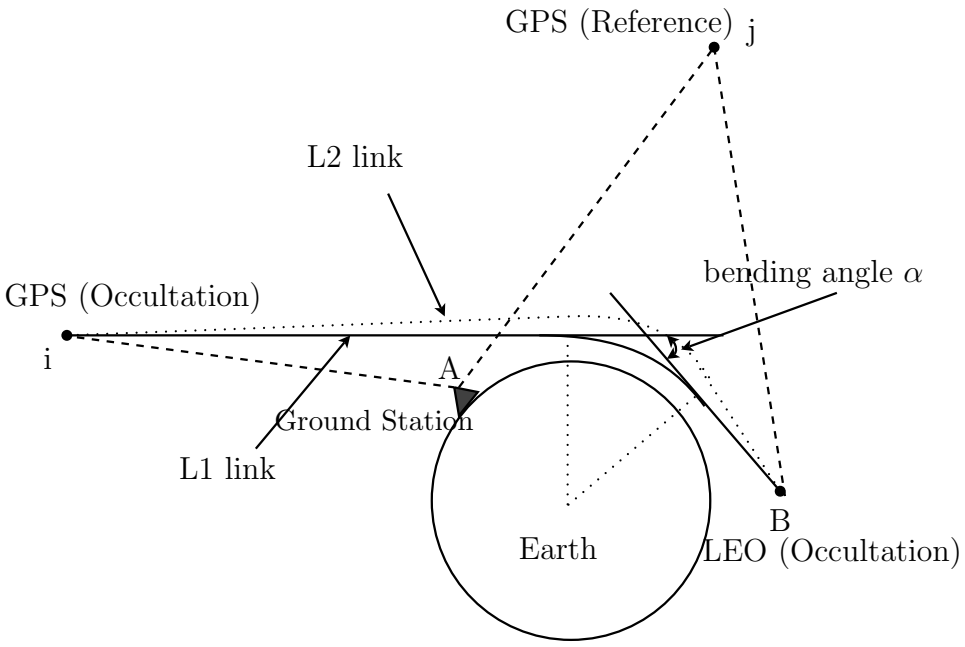


Figure 2.4: The geometric map of single and double differencing.

The GPS phase observation equation between the GPS satellite i and the receiver A is:

$$L_A^i(t)_k = \lambda\Phi_A^i(t) = \rho_A^i(t) + \lambda_k N_A^i(t)_k + c\delta_A(t) + c\delta^i(t) - I_A^i(t)_k + T_A^i(t) \qquad (2.45)$$

Where $L_A^i(t)$ and $\Phi_A^i(t)$ are the measured carrier phases in distance and in cycles respectively, λ is the wavelength, $\rho_A^i(t)$ is the geometric distance between the antenna phase center of satellite i and the receiver at A, c is the speed of light, $N_A^i(t)$ is unknown integer cycles called ambiguity, $\delta_A(t)$ and $\delta^i(t)$ are the receiver and satellite clock biases respectively, I_A^i and T_A^i are phase delays due to the ionosphere and the troposphere along the ray path. The subscript k in the equation indicates the carrier frequency f_1 and f_2. The relativity correction term and the receiver noise term are not included in the above equation, but they will be taken care of in the real data processing.

In order to derive the atmospheric properties, the first measurement we need to obtain is the phase range delay T_B^i due to the existence of the media in the radio link between GPS satellite i and LEO B. In the double differencing technique, besides the GPS occultation satellite and the LEO, another reference GPS satellite and a ground station are required. The GPS receivers at the ground station and onboard the LEO are required to be able to observe both GPS satellites (See Figure 2.4). The single difference $L_A^{ij}(t)$ between the link Ai and Aj is:

$$L_A^{ij}(t) = L_A^i(t) - L_A^j(t) \tag{2.46}$$

The precise orbit of the GPS and LEO satellites can be obtained by adjusting the GPS observations from the ground network and the LEO satellite simultaneously. However, we usually can obtain the GPS precise orbits from the International GPS Service (IGS), calculated by adjusting the measurements from the ground stations in the global fiducial network by data Analysis Centers (AC). The IGS final orbit is a weighted sum of solutions from several ACs. Using the GPS orbit, the Zenith Troposphere delay (ZTD) of the ground stations and the LEO orbits are calculated. We can remove the tropospheric term $T_A^i(t)$ in equation (2.45) from the ZTD of the ground station and a proper mapping function (maps the ZTD to the delay of the real slant path connecting the transmitter and the receiver). The ionospheric effect $I_A^i(t)_k$ can be eliminated in their respective links Ai and Aj through the ionosphere free combination (LC). For link Ai, it is:

$$LC_A^i(t) = \kappa_1 L_A^i(t)_1 - \kappa_2 L_A^i(t)_2 \tag{2.47}$$

with

$$\kappa_1 = \frac{f_1^2}{f_1^2 - f_2^2} \qquad and \qquad \kappa_2 = \frac{f_2^2}{f_1^2 - f_2^2} \tag{2.48}$$

These particular coefficients keep the other terms in the equation the same after the linear combination (except the ambiguity term). Similarly, we can derive the expression for $LC_A^j(t)$. The LC form of the single difference equation (2.46) is:

$$LC_A^{ij}(t) = \rho_A^{ij}(t) + \mu_1 N_A^{ij}(t)_1 - \mu_2 N_A^{ij}(t)_2 + c\delta^{ij}(t) \tag{2.49}$$

with

$$\mu_1 = \frac{cf_1}{f_1^2 - f_2^2} \qquad and \qquad \mu_2 = \frac{cf_2}{f_1^2 - f_2^2} \tag{2.50}$$

For LEO links, the LC equation between the LEO and reference satellite j is equal to:

$$LC_B^j(t) = \rho_B^j(t) + \mu_1 N_B^j(t)_1 - \mu_2 N_B^j(t)_2 + c\delta_B(t) + c\delta^j(t) \tag{2.51}$$

The tropospheric effect term $T_B^j(t)$ is neglected since the link does not penetrate into the troposphere. The double difference form using 3 LC observation links and one frequency dependent occultation link can be given as:

$$LC_{AB}^{ij}(t)_k = LC_A^{ij} - [L_B^i(t)_k - LC_B^j(t)] \tag{2.52}$$

Substitute the terms into the above equation and we obtain:

$$LC_{AB}^{ij}(t)_k = \rho_{AB}^{ij}(t) + I_B^i(t)_k - T_B^i(t) + N \tag{2.53}$$

with

$$N = \mu_1[N_A^{ij}(t)_1 + N_B^j(t)_1] - \mu_2[N_A^{ij}(t)_2 + N_B^j(t)_2] - \lambda_k N_B^i(t)_k \tag{2.54}$$

The excessive tropospheric delay in the occultation link can be derived from equation (2.53):

$$\Delta s_k(t) = T_B^i(t)_k = \rho_{AB}^{ij}(t) + I_B^i(t)_k + N - LC_{AB}^{ij}(t)_k \tag{2.55}$$

This is the stage where the low level data processing ends and the GPS occultation retrieval process starts. According to the equation (2.42), $\Delta s_k(t)$ represents the excessive phase delay in both frequencies.

2.6 The derivation of bending angles from excessive phases

The excessive phase data are first properly filtered using an appropriate running mean to reduce the noise level, and a running standard deviation σ is used to reject the un-reasonable observations (using the 3σ criteria). The Doppler shift in the occultation link is related to the velocity change of the GPS and LEO satellites. It can be derived from the time derivative of the calibrated atmospheric excess phases after appropriate filtering. The derivative of the time series is calculated by the numerical differentiation of a polynomial fit (using Least Squares or SVD) of the excessive phase.

If we rewrite equation (2.42), the total phase change in distance in the path between the GPS and LEO satellites is given as:

$$\Psi(t) = \Psi_0(t) + \Delta\Psi(t) \tag{2.56}$$

The corresponding Doppler frequency change with respect to the carrier frequency is related to the phase change in (2.56) by the relation:

$$(f_d)_k = -\frac{f_k}{c}\frac{d\Psi(t)}{dt} = -\frac{f_k}{c}\dot{\Psi}(t) \tag{2.57}$$

or

$$(\Delta f_d)_k = -\frac{f_k}{c}\frac{d\Delta\Psi(t)}{dt} = -\frac{f_k}{c}\Delta\dot{\Psi}(t) \tag{2.58}$$

$(f_d)_k$ is the total Doppler frequency shift on the frequency k. f_k is the carrier frequency of k. c is the velocity of the light in vacuum. The contribution from $\Psi_0(t)$ causes the frequency shift in vacuum, and the frequency shift $(\Delta f_d)_k$ results from $\Delta\Psi(t)$, which is due to the bending of the ray path caused by the existence of the atmospheric media.

Geometrically from figure 2.3, the change of the optical path $\dot{\Psi}(t)$ is:

$$\dot{\Psi}(t) = \vec{v}_2\vec{m}_2 n_2 - \vec{v}_1\vec{m}_1 n_1 \tag{2.59}$$

where $\vec{v}_1$ and $\vec{v}_2$ represent the velocity vectors of the GPS and LEO respectively, $\vec{m}_1$ and $\vec{m}_2$ are the unit vectors of the signal transmitting direction from the GPS satellite and signal receiving direction to the LEO satellite; and n_1 and n_2 are the refractivities at the satellite locations. From equation (2.57), the Doppler frequency shift has a relationship with the velocity of the GPS and LEO satellites

$$(f_d)_k = -\frac{f_k}{c}(\vec{v}_2\vec{m}_2 n_2 - \vec{v}_1\vec{m}_1 n_1) \tag{2.60}$$

If we define unit vector $\vec{m}_0$ as

$$\vec{m}_0 = \vec{r}/r, \qquad \text{with} \qquad \vec{r} = \vec{r}_2 - \vec{r}_1 \tag{2.61}$$

The frequency shift $(f_{d0})_k$ due to the change of s_0 can be expressed as:

$$(f_{d0})_k = -\frac{f_k}{c}(\vec{v}_2\vec{m}_0 n_2 - \vec{v}_1\vec{m}_0 n_1) \tag{2.62}$$

It is very obvious from Figure 2.3, the bending angle is:

$$\alpha = \phi_1 + \phi_2 + \theta - \pi \tag{2.63}$$

θ can be easily computed from the position vector of the GPS and the LEO satellite. ϕ_1 and ϕ_2 are expected to be obtained by equation (2.60) from the Doppler shift and the velocity of the GPS and LEO satellites. In order to do that, one more equation is needed. If we assume the local spherical distribution of the refractivity

$$n = n(r) \tag{2.64}$$

then the Snell's law equation (2.4) can be expressed in another form:

$$r_1 n_1 \sin\phi_1 = r_2 n_2 \sin\phi_2 \tag{2.65}$$

n_1 and n_2 are unknown in the above equations, since our goal is to derive the refractivity from the occultation measurement, we can not know n_1 and n_2 before these processes. However, we can assume $n_1 = n_2 = 1$, since it can be proved that when using iono-free combination of the bending angles, the effect of these two terms can be neglected [*Gorbunov et al.*, 1996]. Therefore, we can drop these two terms in equation (2.65).

The above equations can be re-formulated in different ways. If we neglect the refractive index at the locations of the satellites, equation (2.59) can be written as:

$$\begin{aligned}
\dot{\Psi} &= v_2^R \cos(\phi_2) + v_2^T \sin(\phi_2) + v_1^R \cos(\phi_1) - v_1^T \sin(\phi_1) \\
&= \dot{r}_2 \cos(\phi_2) + \dot{\theta}_2 r_2 \sin(\phi_2) \dot{r}_1 \cos(\phi_1) - \dot{\theta}_1 r_1 \sin(\phi_2) \\
&= \dot{\theta} p + \frac{\dot{r}_2}{r_2}\sqrt{r_2^2 - p^2} + \frac{\dot{r}_1}{r_1}\sqrt{r_1^2 - p^2}
\end{aligned} \tag{2.66}$$

where r_1 and r_2 are the radial distances of the satellite from the Earth's center. The angle between the two radial vectors is $\theta = \theta_2 - \theta_1$. θ_1 and θ_2 are the angular coordinates of the satellites. The impact factors at two satellite locations are $p = r_1 \sin(\phi_1) = r_2 \sin(\phi_2)$. The $n(r)$ at satellite locations is 1 since satellite orbits are at the outside of the atmosphere. Equation (2.66) will enable us to solve the impact $p(t)$ through the observation $\dot{\Psi}(t)$ and other known time dependent variables. The bending angle has the following geometrical relation with p:

$$\alpha = \theta - \arccos(\frac{p}{r_1}) - \arccos(\frac{p}{r_2}) \tag{2.67}$$

There is no analytical way to solve the non-linear equations (2.60) and (2.65) simultaneously. We can use the iterative method proposed by *Gorbunov et al.* [1996] to solve the problem by first detecting the two angles ϕ_1 and ϕ_2. The initial value of these two angles can be given in vacuum, i.e., ϕ_{10} and ϕ_{20} are the angles between the satellite position vectors and $\vec{m}_0$. It is also defined that $\Delta\phi_k = \phi_k - \phi_{k0}$, and $k = 1, 2$. We have the following iterative steps:

1. Starting from an increment value $\Delta\phi_2$, we compute $\Delta\phi_1$ using the differential form of equation (2.65)

2. The computation of the vector $\vec{m}_{1,2}$ is done by rotating the $\vec{m}_0$ vector at angles $\Delta\phi_{20}$ and $\Delta\phi_{10}$ around the vector normal to the occultation plane.

3. $(\Delta f_d)_k = (f_d)_k - (f_{d0})_k$ is computed using equations (2.60) and (2.62)

4. We compare the $(\Delta f_d)_k$ from the previous step with the computed $(\Delta f_d)_k^{obs}$ in equation (2.58) derived from the observations. If the difference is not small enough, a new value $\Delta\phi_2^{new} = (\Delta f_d)_k^{obs}/(\Delta f_d)_k^{old}\ \Delta\phi_2$ is generated and return to step 1.

We should continue this procedure until it converges. Usually, after several iterations, ϕ_1 and ϕ_2 can be obtained from the converged process. Consequently, the bending angle ϵ is readily derived through equation (2.63).

2.7 The ionosphere correction

Since GPS satellites transmit at two frequencies, the L1 and L2 links between the GPS and LEO satellites travel different paths due to the dispersion property (inverse proportional to f^2) of the ionosphere (see Figure 2.4). The most commonly used method to eliminate the first order ionosphere effect is to apply the so called ionosphere free combination of the phase measurements at two frequencies:

$$L_c(t) = \frac{f_1^2}{f_1^2 - f_2^2} L_1(t) - \frac{f_2^2}{f_1^2 - f_2^2} L_2(t) \tag{2.68}$$

Afterward, the bending angles are derived from the excessive L_c phase delay. However, this method would bring systematic error to the bending angle measurements due to the fact that the two links travel different paths. When combining the phase measurements at the same moment, the impact factor for the two links have different values. Their differences will increase as the altitude goes up. Significant difference may occur at the vicinity of tangent points. It is not feasible to use this combination. To circumvent this problem, *Vorob'ev and Krasil'nikova* [1994] suggested to use the combination of the two bending angles (α_1 and α_2) derived from the two frequency links with the same impact factors.

$$\alpha(p) = \frac{f_1^2}{f_1^2 - f_2^2} \alpha_1(p) - \frac{f_2^2}{f_1^2 - f_2^2} \alpha_2(p) \tag{2.69}$$

The combination effectively removes most of the first order ionosphere effect on the occultation link, which can be seen in Figure 2.5. After the linear combination, LC bending angles are closer to the climatology model MSISE-90 [*Hedin*, 1991] than that of L1 and L2. The remaining residual are mainly due to the non-linearity of the refractivity with respect to the bending angle, but the error is much smaller than the traditional method from equation (2.68), which assumes that the ray trajectories of the two GPS frequencies coincide at the same time t.

2.8 The ellipsoidal correction

In the derivation of bending angles and impact factors, we assume that the atmosphere is spherically symmetric with respect to the Earth's center. However, the

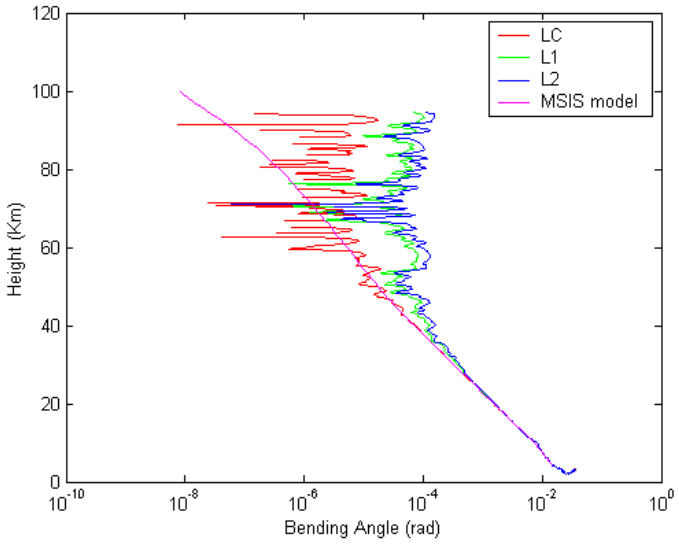


Figure 2.5: The ionospheric correction for bending angles.

more accurate geometric shape of the Earth is not a sphere, but an ellipsoid. The local center of the refractivity does not coincide with the center of the Earth. The study by *Syndergaard* [1998] shows that the neglect of the Earth's oblateness may cause a bias on temperature profile at altitudes below 40 km. The bias is increased with a decrease in altitude and could reach 6°K above the ground for a particular occultation geometry and geographical latitude. The geometry of the occultation configuration with respect to these two centers is shown in Figure 2.6. The movement of the refractive center to the Earth's center will lead to a series of wrong bending angles α^* and impact factors a^*. Consequently, the derived refractivity profile will be wrong. An algorithm was proposed to correct this problem using geometry. We need first to find the location $p(\phi_p, \lambda_p)$ on the ellipsoid where the sphere is tangent to the ellipsoid. The azimuth of the occultation plane (decided by GPS, LEO, and the Earth's center) passing through point p is determined as α_p. The radius of the sphere (radius of curvature–R_{roc}) which is tangent to the ellipsoid at the point p can be calculated by the formula of the geometric geodesy [*Jekeli*, 2000]:

$$R_{roc} = \left(\frac{\cos^2 \alpha_p}{M} + \frac{\sin^2 \alpha_p}{N} \right)^{-1} \tag{2.70}$$

where M and N in above equation are given as:

$$M = \frac{a(1-e^2)}{(1-e^2 \sin^2 \phi_p)^{\frac{3}{2}}} \tag{2.71}$$

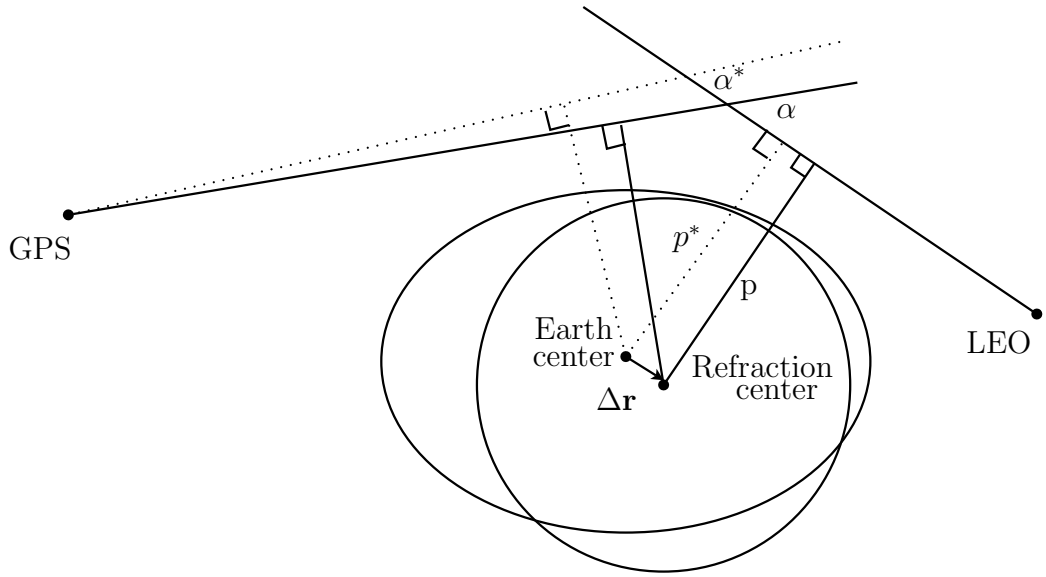


Figure 2.6: The cross section of the oblate Earth and its tangential sphere.

$$N = \frac{a}{1 - e^2 \sin^2 \phi_p} \tag{2.72}$$

M and N are the radius of curvature of the meridian ellipse and the radius of curvature of the prime vertical section at point p, respectively. e is the first eccentricity of the Earth's ellipsoid. Once the radius of curvature is computed, the refractive center departing from the Earth center by $\Delta\mathbf{r}$ can be easily derived using the geometric relationship. All the subsequent profile retrievals will refer to this new center of refractivity.

2.9 Upper boundary limit and statistical optimization

To derive the refractivity profile, we use the inverse Abel transformation represented by equation (2.22). It is noticed that the integral of the equation needs an infinite upper boundary limit. Conventionally, the top level can be defined as the level where no refractivity can be detected by the GPS occultation measurements. Theoretically, as the altitude increases, the bending angle decreases exponentially. However, at the same time, the signal to noise ratio becomes very low. The residual ionosphere error becomes dominant. Generally, at altitudes above 50-60 km, the relative error of the measured bending angle is greater than 100% [*Hocke*, 1997]. The

measurement noise and orbit error contribute to the total error in higher altitude as well.

We rewrite equation (2.22) into two parts, the integration of the first part is from p_0 to p_{top}, an initial impact factor. Its corresponding height is h_{top}. The second part integrates from p_{top} to ∞.

$$n(r_0) = \exp\left[\frac{1}{\pi}\left(\int_{p=p_0}^{p=p_{top}} \frac{\alpha(p)}{\sqrt{p^2 - p_0^2}} dp + \int_{p=p_{top}}^{p=\infty} \frac{\alpha(p)}{\sqrt{p^2 - p_0^2}} dp\right)\right] \tag{2.73}$$

where $p_0 = n(r_0)r_0$ at the tangent point. p_{top} is usually in the range: $R_E + 45$ km $\leq p_{top} \leq R_E + 60$ km. If we choose a top boundary p_{top}, what we will neglect is the contribution from the second integral in equation (2.73). The choice of the initial (top) height can cause significant errors on the temperature profile in the lower stratosphere. In order to retrieve an accurate temperature profile at the stratosphere (e.g., at 40 km), requires a bending angle measurement starting at around 100 km. Unfortunately, the magnitude of the ionospheric component of the measurement is starting to exceed the magnitude of the bending angle from 45-50 km, and continues to dominate the combination above this height. A consistent method should reduce the residual ionospheric error at higher altitude on the final profile retrieval, and at the same time obtain almost the same temperature profile no matter what initial (top boundary) height is chosen.

To minimize the observed bending angle error propagating from the higher altitude to the lower altitude, the bending angles above a certain altitude have to be replaced by some climatology models with small errors. Some typical models include MSISE-90 [*Hedin*, 1991], NRLMSISE-00 [*Picone et al.*, 2002], and CIRA 1986 [*Fleming et al.*, 1988]. *Gorbunov and Sokolovsky* [1993] replace the refraction angle profile with a background model profile above the height where the measurement noise exceeds the normal bending angle variation range. A better way is to derive the new bending angles by combining the climatology model and the observations in an optimized way using statistical information. This was first proposed by *Sokolovskiy and Hunt* [1996], and adopted by many others with improvements and variations, e.g., *Gorbunov et al.* [1996], *Hocke* [1997], *Healy* [2001], etc. The bending angle profile is represented by $\alpha = \alpha(p), p = [bottom, top]$, the two observation equations are:

$$\alpha = \alpha_o + e_o \qquad e_o \sim (0, Q_o) \tag{2.74}$$

$$\alpha = \alpha_m + e_m \qquad e_m \sim (0, Q_m) \tag{2.75}$$

where e_o and e_m are the observation and model error, respectively. Q_o and Q_m are the error covariance matrices, respectively. Solving the above equations using the least square method, we obtain:

$$\alpha_{opt} = (Q_o^{-1} + Q_m^{-1})^{-1}(Q_o^{-1}\alpha_o + Q_m^{-1}\alpha_m) \tag{2.76}$$

Healy [2001] adopted the full model and observation covariance matrices in his study. The difficulty in this method is to find reliable covariance matrices and the assumption of a reasonable correlation height. To simplify the problem, we can neglect the vertical correlation in the observation noise and the model error. The covariance matrices become diagonal, the optimal bending angle for any altitude (or any impact factor) is given as:

$$\begin{aligned}\alpha_{opt}(p) &= \frac{\sigma_m^2(p)}{\sigma_m^2(p)+\sigma_o^2(p)}\alpha_o(p) + \frac{\sigma_o^2(p)}{\sigma_m^2(p)+\sigma_o^2(p)}\alpha_m(p) \\ &= \alpha_m(p) + \frac{\sigma_m^2(p)}{\sigma_m^2(p)+\sigma_o^2(p)}(\alpha_o(p)-\alpha_m(p))\end{aligned} \tag{2.77}$$

In the lower stratosphere, where the noise is not significant, studies show that when comparing the occultation profile with a climatology model (e.g. MSISE-90), the variation of the refractivity is around 5-20%. Thus, an empirical equation for the uncertainty of the model can be written as:

$$\sigma_m(p) = C \cdot \alpha_m(p) \tag{2.78}$$

C is used as a constant for each profile. Most people adopt $C = 0.2$ as a value. The measurement error is estimated by the deviation of the measured bending angle from the modeled bending angle at higher altitude:

$$\sigma_o(p) = \alpha_o(p) - \alpha_m(p) \tag{2.79}$$

This error mainly depends on the ionospheric disturbance. Below the E-layer (A solar controlled ionospheric region around 80-150 km), the error is typically uniform. We can use the value obtained at higher altitude to represent the error below a certain altitude.

A modification of equation (2.77) was given by *Hocke* [1997]. He replaced the σ^2 with σ in the equation:

$$\alpha_{opt}(p) = \alpha_m(p) + \frac{|\sigma_m(p)|}{|\sigma_m(p)|+|\sigma_o(p)|}(\alpha_o(p)-\alpha_m(p)) \tag{2.80}$$

This modification will change the shape of the weighting function.

2.10 The numerical solution of the bending angle

The refractivity profile is solved by the inverse Abel transform equation (2.22). The integral has a singularity at $p = p_0$. To avoid this "pole" problem, an alternative form of the inverse Abel transformation equation is derived from the original form. Considering the relationship:

$$\int \frac{dp}{\sqrt{p^2-p_0^2}} = \ln\left(|p| + \sqrt{p^2-p_0^2}\right) \tag{2.81}$$

and starting from equation (2.22), we integrate the right hand side by parts:

$$n(r_0) = \exp\left[\frac{1}{\pi}\left(\alpha(p)\ln\left(p+\sqrt{p^2-p_0^2}\right)\Big|_{p=p_0}^{p=\infty} - \int_{p=p_0}^{p=\infty}\ln\left(p+\sqrt{p^2-p_0^2}\right)d\alpha\right)\right] \quad (2.82)$$

We change the expression $\ln\left(p+\sqrt{p^2-p_0^2}\right)$ to another form:

$$\ln\left(p+\sqrt{p^2-p_0^2}\right) = \ln p_0 + \ln\left(\frac{p}{p_0}+\sqrt{\frac{p^2}{p_0^2}-1}\right) \quad (2.83)$$

Substituting into equation (2.82), it finally becomes:

$$n(r_0) = \exp\left[-\frac{1}{\pi}\int_{\alpha=\alpha(p_0)}^{\alpha=0}\ln\left(\frac{p}{p_0}+\sqrt{\frac{p^2}{p_0^2}-1}\right)d\alpha\right] \quad (2.84)$$

This expression avoids the problem that occurs in equation (2.22), and is suitable for numerical integration.

Besides the direct numerical integration method, one also can invert the forward Abelian integration equation (2.16). It can be expressed in discrete form:

$$\frac{\alpha_m}{2p_m} = -\sum_{k=1}^{m}\nabla\tilde{n}_k\frac{\Delta x_k}{\sqrt{x_k^2-p_m^2}} \quad (2.85)$$

In the above equation, the atmosphere is divided into m layers with the thickness of Δx_k. $\nabla\tilde{n}_k$ is the gradient of the refractive index in layer k, which is assumed to be constant within the layer. The matrix form of the above equation (2.85) is:

$$\left(\frac{\vec{\alpha}}{2p}\right) = \mathbf{A}\nabla\vec{\tilde{n}} \quad (2.86)$$

or

$$\begin{pmatrix} \frac{\alpha_1}{2p_1} \\ \frac{\alpha_2}{2p_2} \\ \frac{\alpha_3}{2p_3} \\ \vdots \\ \frac{\alpha_m}{2p_m} \end{pmatrix} = \begin{pmatrix} A_{11} & 0 & 0 & \dots & 0 \\ A_{21} & A_{22} & 0 & \dots & 0 \\ A_{31} & A_{32} & A_{33} & \dots & 0 \\ \vdots & & & & \\ A_{m1} & A_{m2} & A_{m3} & \dots & A_{mk} \end{pmatrix} \begin{pmatrix} \nabla\tilde{n}_1 \\ \nabla\tilde{n}_2 \\ \nabla\tilde{n}_3 \\ \vdots \\ \nabla\tilde{n}_k \end{pmatrix} \quad (2.87)$$

With the help of equation (2.81), we can re-write A_{mk} to avoid problems when $p = p_0$:

$$A_{mk} = -\frac{\Delta x_k}{\sqrt{x_k^2-p_m^2}} = -\ln\left(|x_k|+\sqrt{x_k^2-p_m^2}\right)+\ln\left(|x_{k-1}|+\sqrt{x_{k-1}^2-p_m^2}\right) \quad (2.88)$$

The least square solution of equation (2.86) is:

$$\nabla\vec{n} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\left(\frac{\vec{\alpha}}{2p}\right) \tag{2.89}$$

We can accumulate the layered refractive index gradient to obtain the refractive index profile. Equation (2.87) can also be solved using the onion peeling method, which is to divide the atmosphere into thin layers and reconstruct the homogeneous refractivity in each layer sequentially, like peeling an onion. This is a classical way of performing ray tracing. Using the characteristics of the triangle matrix A, we can solve the refractivity layer by layer from the top. The weakness of this method is the propagation of the error from the top layer to the bottom layer, since the inversion of the subsequent layers will depend on the results from the previous layers.

A general procedure of retrieving a GPS occultation profile using geometric optical method is summarized in Figure 2.7.

2.11 The spatial resolution of GPS occultation

In the previous sections, we derived the atmospheric refractive index based on the assumption of the geometric optical ray trajectory, i.e., the GPS signal passing through the Earth's atmosphere and arriving on the receiver onboard the LEO follows a dimensionless ray trajectory, which assumes no transverse dimension on the cross-section of the ray path. However, the geometrical ray trajectory is only an approximation to the Fresnel diffraction as the wavelength approaches zero.

2.11.1 Vertical resolution

According to the Huygens-Fresnel theory [*Born and Wolf*, 1993], the electromagnetic (EM) field at the receiver point equals the superposition of the secondary waves that proceed from a surface situated between this point and the transmission source. The theory leads to a cross-section of the beam at the limb consisting of concentric Fresnel zones with interlaced positive and negative contributions centered at the ray trajectory. Since the contribution from adjacent Fresnel zones cancels, the sampling can be effectively represented by the first Fresnel zone [*Born and Wolf*, 1993], which can be characterized as the "thickness"of the geometrical optical ray path. Therefore, the atmospheric delay sensed by the receiver will dominantly be affected by the refractivity inside the finite Fresnel volume. The first Fresnel zone is defined as the region where the path delay inside the region does not exceed the minimum path delay of the region by half of the wavelength (see Figure 2.8). To determine the radius of the first Fresnel zone r_f, the following relationship must be fulfilled:

$$(L_1 + L_2) - (l_1 + l_2) = \frac{\lambda}{2} \tag{2.90}$$

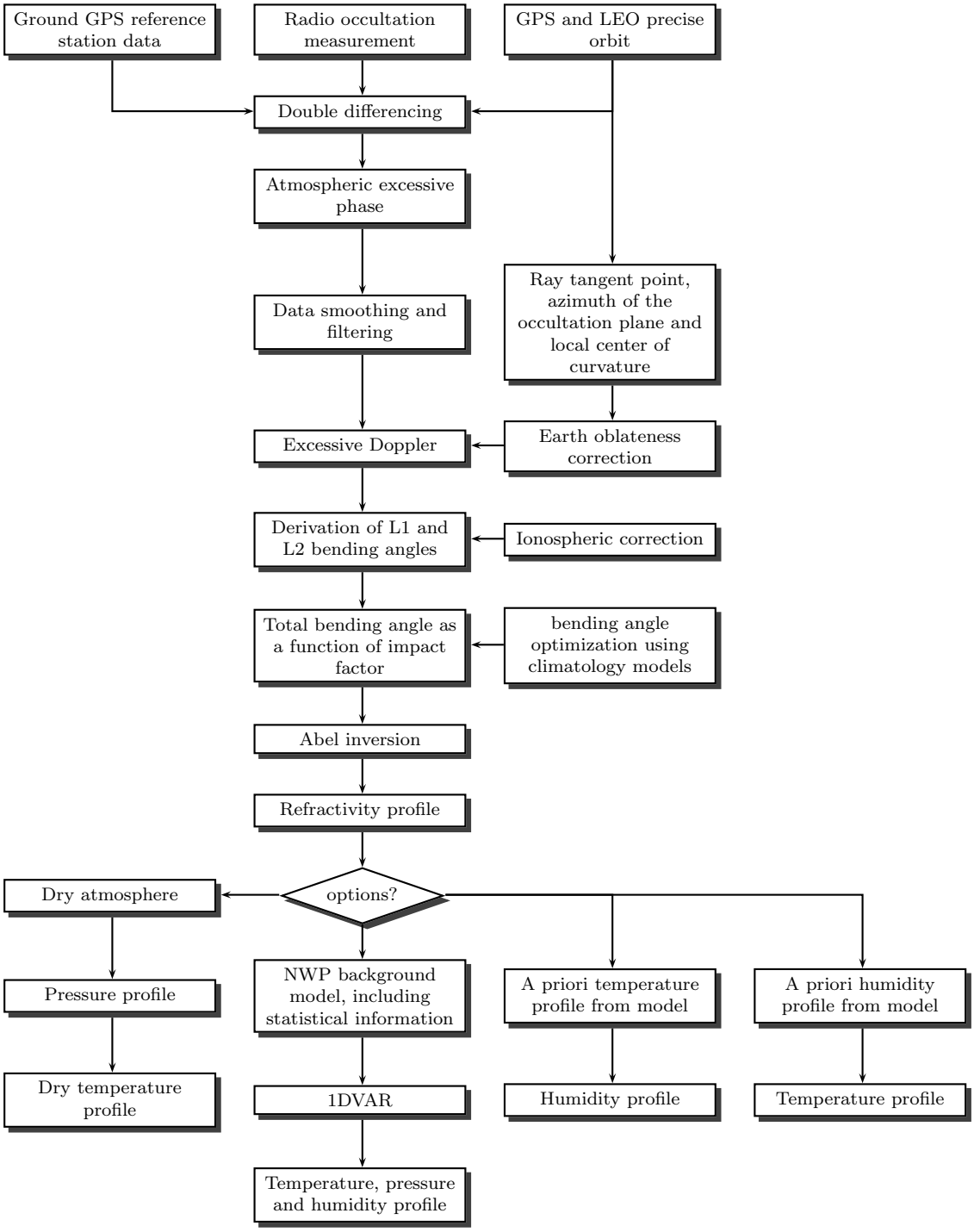


Figure 2.7: The block diagram of radio occultation data processing.

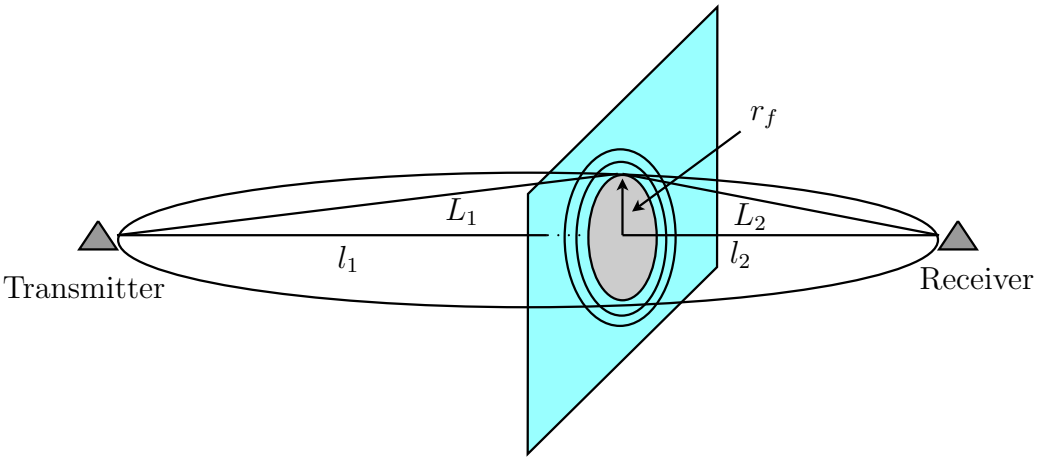


Figure 2.8: The cross-section of the first Fresnel zone.

with

$$L_1 = \sqrt{l_1^2 + r_f^2} \qquad and \qquad L_2 = \sqrt{l_2^2 + r_f^2} \tag{2.91}$$

Solving for r_f from equation (2.90), we obtain:

$$r_f = \sqrt{\frac{\lambda l_1 l_2}{l_1 + l_2}} \approx \sqrt{\lambda l_2} \qquad (if\ l_1 \gg l_2) \tag{2.92}$$

For a general configuration of GPS satellites and the LEO satellite, l_1 usually is around 25000 km, and l_2 is around 3000 km. The assumption of $l_1 \gg l_2$ is valid. Equation (2.92) gives us the first Fresnel zone in a non-refracting medium. In a refracting medium, the first Fresnel zone also depends on the gradient of the refractivity. Its radius in the presence of the refractivity is given approximately as

$$r_f(p) = \sqrt{\lambda l_2 \zeta(p)} \tag{2.93}$$

where $\zeta(p)$ is the defocusing factor and is given by [*Melbourne et al.*, 1994]:

$$\zeta(p)^{-1} = 1 - \frac{l_1 l_2}{l_1 + l_2} \frac{d\alpha}{dp} \tag{2.94}$$

Theoretically, as the refractivity increases, the Fresnel zone decreases and the resolution is improved as it approaches to the lower atmosphere. For the L1 carrier of the GPS, the wavelength is around 19 cm, and equation (2.92) yields a value for the first Fresnel zone radius of 750 m in the upper stratosphere to around 275 m in the

lower troposphere. Since the L1 and L2 carriers of the GPS satellite have different wavelengths, the radii of the first Fresnel zones differ by 13%. The diameter of the first Fresnel zone gives an approximate vertical resolution around 1.5 km in the upper atmosphere.

2.11.2 Horizontal resolution

The horizontal resolution can be approximately given out as the length of the ray path that is tangent to two concentric circles whose radii differ by a vertical resolution of Δh. We simplify the calculation using Figure 2.9. The horizontal and vertical resolutions are related by the expression:

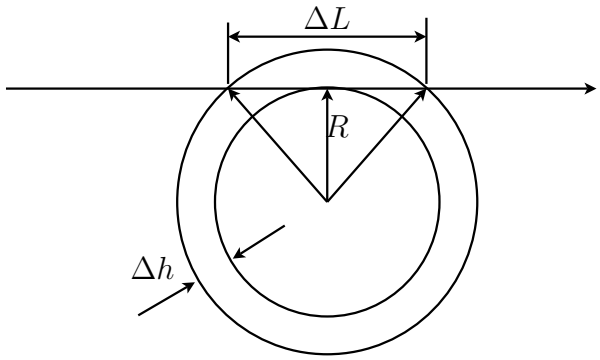


Figure 2.9: The horizontal resolution along the ray trajectory.

$$\Delta L \approx 2\sqrt{2R\Delta h} \tag{2.95}$$

where R is the radius of the atmosphere from the ray tangent point. Given a vertical resolution of 1.5 km, the horizontal resolution is about 277 km. Therefore, the bending angle α measured at the ray perigee is an integral effect, which is contributed to from the refractivity inside a tube along the ray path with the length of ΔL. The distribution of the contribution to the bending angle along the ray path is like a Gaussian distribution, with the largest contribution near the ray perigee point.

CHAPTER 3

ADVANCED METHODS FOR GPS OCCULTATION IN MULTIPATH REGIONS

3.1 Atmospheric multipath

In GPS occultation, atmospheric multipath happens when multiple rays arrive at the receiver on LEO at the same time (Figure 3.1). What the receiver measures is the superposition of the wave signals from different directions. The way of retrieving the bending angle from the standard "geometrical optics" method is based on the fact that the spatial gradient of the phase of the electromagnetic wave along the satellite trajectory is uniquely related to the direction of the ray projected onto the trajectory. Only one ray comes to the receiver at a time, thus, every bending angle in the profile should have a single corresponding impact factor, i.e. the bending angle is a monotonous function of the impact factor and the ray paths are separated. In the previous chapter, a GPS signal was assumed to propagate as a dimensionless ray. This is valid when the wavelength of the signal is small and the vertical gradient of the refractive index is within a certain limit. However, in the lower troposphere, the rapid change of the moisture content increases the chance of the formation of sharp refraction layers. The strong vertical refractive index gradient of these layers will cause multipath propagation of the signal and the diffraction effect becomes significant. In this case, the direct and reflected rays arrive at the receiver antenna at the same time. The phase measured is the superposition of the phase and amplitude of the interfering rays (see Equation (3.4) on page 39). This interfering phase measurement can not be used to derive the refractive angle, since in this region, the geometric optics method would result in multiple correspondences of bending angles to impact factors. In order to use the Abel inversion approach, we have to select a single pair of bending angle and impact factor to avoid an ambiguity in the profile. The multipath leads to substantial errors in the results. The standard "geometrical optics" method is incapable of separating the ray paths in the multipath region. Furthermore, the complicated, strong vertical gradient structure causes significant signal tracking problem for GPS occultation receivers. Under this circumstance, the tracking quality will be degraded with larger errors and the loss of lock will frequently happen. We can derive the condition

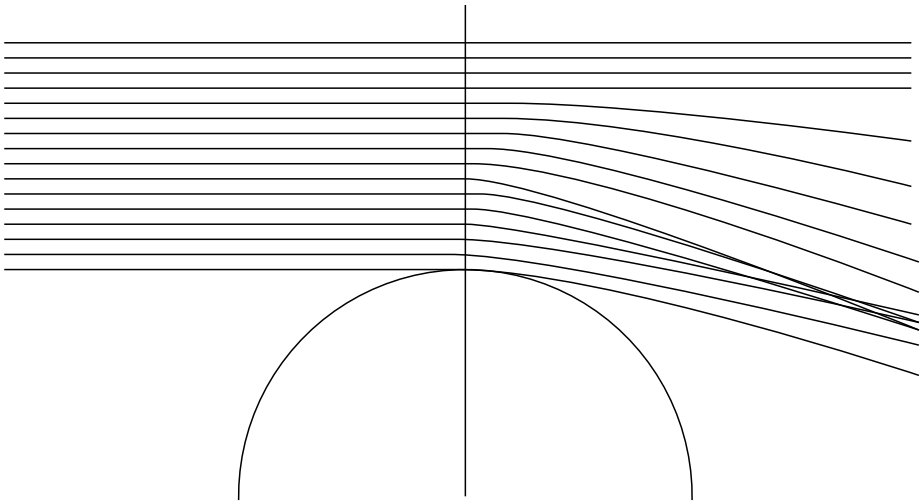


Figure 3.1: The illustration of Multipath propagation of rays.

for multi-ray crossing at the receiver with an auxiliary Figure 3.2. The solid curve is an ideal atmospheric relation between the bending angle and the impact factor. $\alpha(p)$ is a monotonic decreasing function. The dashed line shows a relationship between the bending angle and the impact factor at a specific measurement epoch by equation (2.67). The derivative of equation (2.67) with respect to p (for fixed GPS and LEO position) is:

$$\frac{d\alpha_t(p)}{dp} = \frac{1}{L_1} + \frac{1}{L_2} \qquad \text{with} \qquad L_i = \sqrt{L_i^2 - p^2}, i = 1, 2 \tag{3.1}$$

It is a monotonic increasing function. For a single-ray propagation, the two curves $\alpha(p)$ and $\alpha_t(p)$ only have one intersection point. However, if

$$\frac{d\alpha(p)}{dp} > \frac{1}{L_1} + \frac{1}{L_2} = \frac{L_1 + L_2}{L_1 L_2} \approx \frac{1}{L_2} \tag{3.2}$$

there are multiple pairs of bending angles and impact factors corresponding to this measurement at a specific time, i.e., the choice of rays is not unique.

3.2 Single-ray propagation vs. multi-ray propagation

Single ray propagation refers to the situation that only one ray arrives at every observation point along the orbit. In this situation, the arriving angle of the ray can be determined by the derivative of the phase. Figure 3.3 illustrates how a single

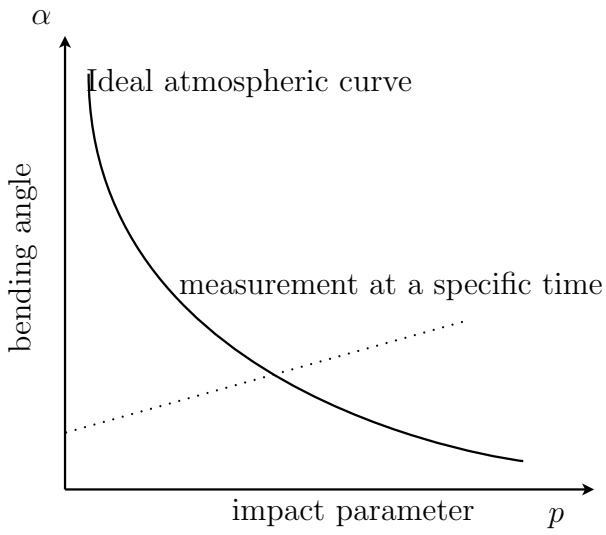


Figure 3.2: The derivation of the condition of multi-ray intersection.

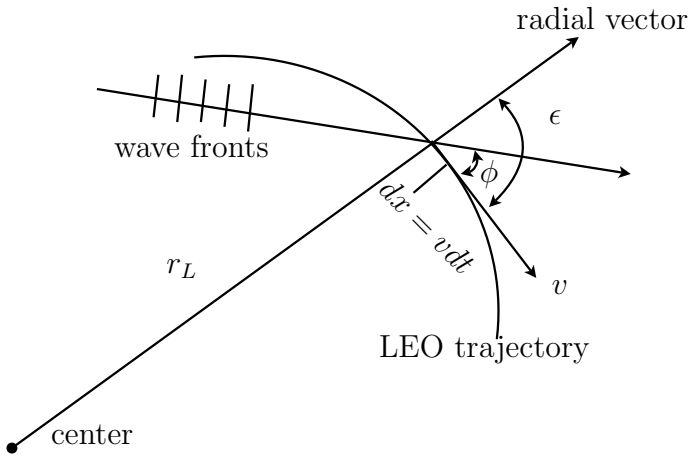


Figure 3.3: The single ray propagation.

ray propagation allows the determination of the direction of the ray. We assume the transmitter is stationary and the atmosphere is symmetric. The ray equation is described by $u(x) = A(x)\exp(i\Phi(x))$. In the occultation plane, only the motion

across the wave fronts contributes to the phase measurements:

$$\frac{d\Phi}{dx} = k\cos\phi \qquad \text{or} \qquad f_d = \frac{d\Phi}{dt} = kv\cos\phi \tag{3.3}$$

where $k = \frac{2\pi}{\lambda}$, is the wave number, ϕ is the angle between the satellite trajectory and the ray direction. However, when more than one ray arrive at one location, as shown

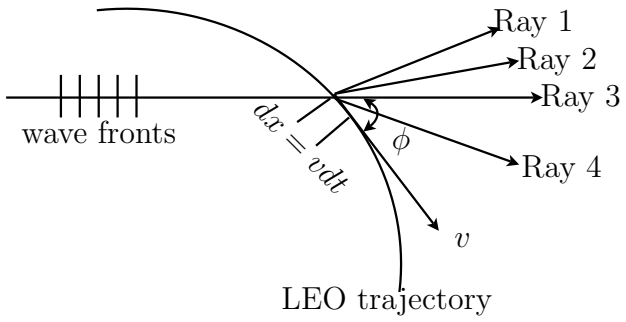


Figure 3.4: Multiple rays propagation.

in Figure 3.4, the measured complex wave field is:

$$u(x) = A(x)\exp(i\Phi(x)) = \sum_{i=0}^{i=N} A_i(x)\exp(i\Phi_i(x)) \tag{3.4}$$

N is the total number of rays at the location. From the equation above, it is hard to determine any of the ray directions from the derivative of the total phase.

3.3 Radio holographic methods

Radio holographic methods are a class of methods which are based on the analysis of the received complex electromagnetic signal (amplitude and phase) from the receiver trajectory.

3.3.1 Radio-optic (slide spectral) method

In single ray areas, the incident ray direction can be inferred from differentiating the observed signal phase with respect to the time. As shown above, if multiple rays reach one location, the derivative of the total phase at every epoch does not automatically give us the instantaneous frequency for each ray. The frequencies can

be identified by analyzing the local spectral of the complex signal within a small aperture (sliding window). The local maxima of the spectral amplitude correspond to different rays arriving at the center of the aperture. The frequency of each spectral maxima defines the ray arriving angle. The arriving angle along with the position of the aperture center determine the bending angles and impact factors of the rays arriving in the aperture. This method can be considered as a trade-off between the time resolution and frequency resolution, i.e., the aperture must be large enough for a suffice spectral resolution and must be small enough for a better resolution of impact factor. The radio-optic method was originally used for planetary atmosphere [*Lindal et al.*, 1987], and was later introduced to the Earth's atmosphere [*Hocke et al.*, 1999; *Gorbunov et al.*, 2000; *Sokolovskiy*, 2001].

The Fourier spectral of the sliding window is represented as $F(f)$, where the frequency f is associated to the ray arriving angle. From equation (3.3), the angle between the ray directory and satellite trajectory is:

$$\phi = \arccos(\frac{f}{kv}) \tag{3.5}$$

Geometrically from Figure 3.3, the impact factor is:

$$p = r_L \sin(\epsilon - \phi) \tag{3.6}$$

where ϵ is the angle between the radial vector and LEO velocity vector. The bending angle is obtained from equation (2.63):

$$\alpha = \arcsin(\frac{p}{r_G}) + \arcsin(\frac{p}{r_L}) + \theta - \pi \tag{3.7}$$

This method is theoretically simple, but it is often difficult to automatically identify the local maxima when the ray structure is complicated. *Sokolovskiy* [2001] proposed a method that uses the spectral power to weigh the spectral content inside the aperture. This method can well resolve the bending angle as a function of the impact factor, with no need to find the local maxima. In the real data processing, a reference complex signal $u_m(t)$ calculated from a smooth model of the refractivity is required to correct the wave front curvature for the focus of the wave to the aperture [*Igarashi et al.*, 2000]. This method only has good resolution for the bending angle and impact factor in the lowest troposphere where strong refraction exists, and sometimes the diffraction may result in wrong maxima in subcaustic zones [*Gorbunov et al.*, 2000].

3.3.2 The back propagation method (BP)

In the back propagation method, the complex field (amplitude and phase) at LEO is propagated back to a plane in the atmosphere, where there is no multipath and rays are well separated. Thus, the standard geometrical optics method can be used

in this region to reconstruct the profile with improved resolution and less multipath effect.

The wave propagation in the atmosphere can be described in terms of a scalar wave equation in the inhomogeneous media [*Chew*, 1990]:

$$(\Delta + k^2 n^2) u(\mathbf{r}) = 0 \tag{3.8}$$

where $k = \frac{2\pi}{\lambda}$ is the wavenumber, n is the refractivity, Δ is the Laplacian, and u is the complex wave field. The above equation is also called Helmholtz equation. Being practical, we only use the 2D form of the Helmholtz equation in the GPS radio occultation plane by neglecting the diffraction effect in the lateral direction. This can be justified by the lateral size of the first Fresnel zone $\sqrt{L\lambda}$. L and λ are the distances from the receiver to the ray perigee and the wavelength of the signal respectively. The typical value for L is 3000 km and λ is 20 cm. Consequently, the first Fresnel zone in the lateral direction is less than 1 km, which is significantly smaller than the nominal atmospheric inhomogeneities in the lateral direction. This allows us to neglect the diffraction effects in this dimension. On the other hand, however, the vertical size of the first Fresnel zone is similar with the vertical atmospheric features, and it cannot be simply neglected.

We denote the position vector of the source and observation vector as $\mathbf{r}$ and $\mathbf{r}'$ respectively. The Green function of the 2-D Helmholtz equation is given as [*Chew*, 1990]:

$$G(\mathbf{r}, \mathbf{r}') = \frac{1}{4i} H_0^{(1)}(k\rho) \qquad with \qquad \rho = |\mathbf{r} - \mathbf{r}'| \tag{3.9}$$

where $H_0^{(1)}(k\rho)$ is the Hankel function of the first kind and 0th order. $\mathbf{r}$ and $\mathbf{r}'$ are the vectors in the 2D occultation plane.

If the complex field along a curve boundary S (LEO orbital trajectory) is known as $u(\mathbf{r}')$, the field at an arbitrary position $\mathbf{r}$ can be given by the solution of the Helmholtz equation, which is a 2D external boundary problem:

$$u(\mathbf{r}) = \frac{i}{2} \int_S u(\mathbf{r}') \frac{\partial}{\partial n_{r'}} H_0^{(1)}(k|\mathbf{r} - \mathbf{r}'|) dS_{r'} \tag{3.10}$$

where $n_{r'}$ is the external normal to the boundary S. For large argument $|\mathbf{r} - \mathbf{r}'| \to \infty$, the Hankel function in equation (3.9) can be approximated by:

$$H_0^{(1)}(k|\mathbf{r} - \mathbf{r}'|) = \sqrt{\frac{2}{\pi k |\mathbf{r} - \mathbf{r}'|}} \exp[i(k|\mathbf{r} - \mathbf{r}'| - \frac{\pi}{4})] \tag{3.11}$$

Substituting equation (3.11) into equation (3.10), the equation for the wave to propagate from the curves source S is given as:

$$u(\mathbf{r}) = \sqrt{\frac{k}{2\pi}} \int_S u(\mathbf{r}') \Big[\frac{\exp(ik|\mathbf{r} - \mathbf{r}'| - \frac{\pi}{4})}{|\mathbf{r} - \mathbf{r}'|^{3/2}} |\mathbf{r} - \mathbf{r}'| \cdot n_{r'}) \Big] dS_{r'} \tag{3.12}$$

Similarly, the equation for the reconstruction of the source field is:

$$u(\mathbf{r}) = \sqrt{\frac{k}{2\pi}} \int_S u(\mathbf{r}') \Big[\frac{\exp(ik|\mathbf{r}-\mathbf{r}'| + \frac{\pi}{4})}{|\mathbf{r}-\mathbf{r}'|^{3/2}} |\mathbf{r}-\mathbf{r}'| \cdot n_{r'}) \Big] dS_{r'} \tag{3.13}$$

Equations (3.12) and (3.13) are two reciprocal transformations.

In most of the situations, the occultation geometry does not lie in a single plane. Through introducing a proper coordinate system and a suitable transformation method, the GPS satellite can be treated as stationary and the LEO orbital trajectory can be approximately treated as a line during the occultation event. This transformation will conserve the resulted bending angles and impact factors of the rays. We still can use the regular geometrical optics method to calculate the profiles.

The back propagation plane must be placed inside the atmosphere as well as in a single-ray zone. Without any prior knowledge of the atmosphere, it is usually not easy to decide where to place the back propagation plane. *Gorbunov et al.* [2000] suggested an empirical formula to calculate the plane position as $x = R_E \alpha_{max}$. R_E is the radius of the curvature of the Earth and α_{max} is the maximum bending angle in the occultation. This position is located between the real and imaginary caustics. We can also use $x = 200$ km as a typically suggested value.

3.3.3 Canonical transform method (CT)

The back propagation method requires the choice of an appropriate back propagation plane in the single ray area. In the complicated situation such that the real and imaginary caustics overlap, the single ray area does not exist or is very hard to locate. This will cause the back propagation method to fail. In order to deal with this difficulty, a new approach was introduced by *Gorbunov* [2002] based on the theory of Fourier Integral Operator [*Egorov et al.*, 1999] associated with the canonical transformation [*Arnold*, 1978]. The complex field $u(x, y)$ measured along the LEO trajectory in the occultation plane is first needed to be back-propagated to a vertical line at a position x. u now becomes a function of the vertical coordinate y. The geometrical optics can be represented by a Hamilton system in the phase space with coordinate y and momentum η. This phase space coordinate (y, η) is required by the canonical method. In a single ray area, the momentum can be explained as the derivative of the eikonal of the wave field (or ray direction vector). If we find another coordinate (p, ξ) representing the same phase space, the rays can be described using the new coordinates. In the new phase space, the new momentum ξ is a single-value function of the new coordinate p. This allows us to disentangle the interfering rays, since they always have different momenta. p is typically the ray impact factor, and the new momentum ξ is the ray direction angle, which can be used to derive the bending angle using simple geometric relationships. The transformation from the geometrical coordinate (y, η) to the ray coordinate (p, ξ) is canonical and the phase

space is preserved. The transformation can be implemented asymptotically using the Fourier integral operator.

In this section, we briefly introduce the canonical method [*Gorbunov*, 2001, 2002] for an asymptotic solution of the wave equation using the Fourier integral operator [*Egorov et al.*, 1999; *Mishchenko et al.*, 1990].

We use Cartesian coordinates (x, y) to describe the wave field $u(x, y)$ in the occultation plane. x points to the wave propagation direction, and y is perpendicular to x. The wave field measured along the LEO trajectory needs to be back propagated to a vertical line at an arbitrary location x. It can be denoted by $u_x(y)$. We introduce the normalized differential operators $D_x = \frac{1}{ik}\frac{\partial}{\partial x}$ and $D_y = \frac{1}{ik}\frac{\partial}{\partial y}$. The Helmholtz equation (3.8) in vacuum becomes:

$$(-D_x^2 - D_y^2 + 1)u = 0 \tag{3.14}$$

For very short waves, the backscattering can be neglected. The above equation can be factorized as [*Martin*, 1992]:

$$(D_x + \sqrt{1 - D_y^2})(-D_x + \sqrt{1 - D_y^2})u = 0 \tag{3.15}$$

Then, the pseudo-differential equation describes the wave propagating forward in the x direction is:

$$-D_x u = H(y, D_y) \qquad \text{with} \qquad H(y, D_y) = -\sqrt{1 - D_y^2}u \tag{3.16}$$

This is the 2D form of the "generalized parabolic equation" derived in *Martin* [1992]. $H(y, D_y)$ is the Hamilton operator. To derive the asymptotic solution of equation (3.16), the wave field equation is expressed in the form $u(x, y) = A(x, y)\exp(ik\Psi(x, y))$. We substitute this into equation (3.16) and equate the coefficients according to the same power in k^{-1}. We also neglect the derivative of the amplitude, since it only contributes to the coefficients of the next order. The coefficients of the highest order (k^0) yield the Hamilton-Jacobi equation:

$$-\frac{\partial \Psi}{\partial x} = H(y, \eta), \qquad \text{with} \qquad \eta = \frac{\partial \Psi}{\partial y} \tag{3.17}$$

The above equation tells us that the unit ray direction vector is $\nabla\Psi = (\sqrt{1 - \eta^2}, \eta)$, and η geometrically represents the vertical projection of the ray direction vector.

The Hamilton-Jacobi equation can be reduced to a characteristic set of ordinary differential equations:

$$\dot{y} = \frac{dy}{dx} = \frac{\partial H}{\partial \eta}, \qquad \dot{\eta} = \frac{d\eta}{dx} = \frac{\partial H}{\partial y}, \qquad \dot{\Psi} = \frac{d\Psi(x, y)}{dx} = \eta\dot{y} - H \tag{3.18}$$

Given a boundary condition in the source plane $u_0(y) = A_0(y)\exp(ik\Psi_0(y))$, without knowing the ray structure, we do not know the relationship between the wave field and

the ray direction. For single ray propagation, the momentum $\eta(y)$ at points along the vertical line can be calculated by $\frac{\partial \Psi_0(y)}{\partial y}$. Using this initial condition, we can integrate equation (3.18) to find the ray destination at other locations. For an arbitrary ray structure, however, the relationship in equation (3.17) cannot be applied. Moreover, when the multipath occurs, the geometrical optics amplitude is singular at the vicinity of the caustics.

To find a general asymptotic solution to the Cauchy problem in the refracting medium with the existence of the caustics, we shall use Egorov's theorem [*Egorov et al.*, 1999; *Mishchenko et al.*, 1990], which states that the Fourier Integral Operator (FIO) is associated with a canonical transformation from the old coordinates to the new coordinates in the phase space. The Hamilton system of the geometric optical ray describes the dynamics of the ray by coordinates (y, η) in the phase space. We have options to choose different canonical coordinates (p, ξ) in the same phase space to represent the same Hamilton system. In a spherical symmetric atmosphere, the impact factor p can specify a unique ray. Thus, we introduce the new coordinates (p, ξ) in the phase space, which can disentangle the multiple rays with different momenta. Along with the change of the coordinates, the wave field can also be asymptotically transformed to the new coordinates using Egorov's theorem. For the old coordinates (y, η) in the phase space, the Hamilton system is in the form $-D_x u = H(y, D_y)u$. After introducing the new coordinates (p, ξ), the new Hamilton function becomes $H'(p, D_p)$. The wave field in the new coordinates is u', which is asymptotically equal to $\hat{\Phi}u$, where $\hat{\Phi}$ is the Fourier integral operator. In this way, we write the new Hamilton system as $-D_x u' = H'(p, D_p)u'$ with the initial condition $\hat{\Phi}u_0$. The momentum ξ in the new coordinates is the ray direction angle, which can be obtained by taking the derivative of the phase of the transformed wave field with respect to the impact factor p. Finally, the ray direction ξ can be linked to the bending angle by geometric relationships.

We first define the 1D k-Fourier transformation like in *Gorbunov* [2002]:

$$\tilde{u}(\eta) = \int e^{-iky\eta} u(y) dy \tag{3.19}$$

$$u(y) = \frac{k}{2\pi} \int e^{iky\eta} \tilde{u}(\eta) d\eta \tag{3.20}$$

The Fourier integral operator can be given in the following form [*Mishchenko et al.*, 1990]:

$$u_x(p) = \frac{k}{2\pi} \int a(p, \eta) e^{ik(\Sigma(p,\eta)+y\eta)} \tilde{u}_0(\eta) d\eta \tag{3.21}$$

where y is the starting point of the single ray with momentum η and at the end point p, i.e. $y = y(p, \eta)$. Functions $a(p, \eta)$ and $\Sigma(p, \eta)$ describe the amplitude and phase delay of the ray. Since the ray direction is perpendicular to the ray fronts, we have the differential equation:

$$d\Sigma = \xi dp - \eta dy \tag{3.22}$$

We define $S(p,\eta) = \Sigma(p,\eta) + y\eta$, then

$$dS = \xi dp + y d\eta \tag{3.23}$$

S is the generating function of the canonical transformation [*Arnold*, 1978]. Given a generating function S, the canonical transform is defined as:

$$y = \frac{\partial S(p,\eta)}{\partial \eta}, \qquad \xi = \frac{\partial S(p,\eta)}{\partial p} \tag{3.24}$$

We choose the new coordinate p to be equal to the impact factor to distinguish rays. The impact factor is the distance from $(0,0)$ to the ray, which can be derived by taking the dot product of a point (x,y) and the unit vector perpendicular to the ray direction vector $(\sqrt{1-\eta^2}, \eta)$:

$$p = -x\eta + y\sqrt{1-\eta^2} \tag{3.25}$$

then, the generating function becomes:

$$S(p,\eta) = \int y(p,\eta) d\eta = \int \frac{p + x\eta}{\sqrt{1-\eta^2}} = p \arcsin\eta - x\sqrt{1-\eta^2} \tag{3.26}$$

Using equation (3.24), the momentum ξ in the new coordinates is

$$\xi = \arcsin\eta \tag{3.27}$$

The amplitude of the ray at the destination point z is derived using the conservation of the energy flux in the ray tube [*Kravtsov and Orlov*, 1990]:

$$A(p) = A_0(y)\sqrt{\frac{dy}{dp(y,\eta(y))}} \tag{3.28}$$

Function $a(p,\eta)$ in equation (3.21) reflects the variation of the amplitude. Using equation (3.28) along with equations (3.24) and (3.26), the amplitude $a(p,\eta)$ is derived as:

$$a(p,\eta) = \sqrt{\frac{dy}{dp}} = \sqrt{\frac{\partial^2 S}{\partial p \partial \eta}} = (1-\eta^2)^{-\frac{1}{4}} \tag{3.29}$$

By substituting equations (3.26) and (3.29) into equation (3.21), and the Fourier integral operator takes the following form:

$$u_x(p) = \hat{\Phi} u_x = \frac{k}{2\pi} \int (1-\eta^2)^{-\frac{1}{4}} e^{ik(p\arcsin\eta - x\sqrt{1-\eta^2})} \tilde{u}_x(\eta) d\eta \tag{3.30}$$

In this way, the wave field is transformed to the expression in the form of the impact factor p by the operator equation (3.30) and generating function (3.26). If we express

transformed field as $u(p) = A'(p)\exp(ik\Psi'(p))$, the momentum, or the direction angle ξ is equal to $\frac{\partial\Psi'}{\partial p}$. Finally, the bending angle is computed as:

$$\alpha(p) = -\xi(p) + \arcsin\left(\frac{|x_{GPS}| - \sqrt{x_{GPS}^2 - 2a(p-a)}}{a}\right) \tag{3.31}$$

where a is the earth radius, and the second term in the equation corrects the ray direction angle at the GPS satellite location (x_{GPS}, y_{GPS}).

3.3.4 Principle of phase matching and full spectrum inversion method (FSI)

3.3.4.1 Principle of phase matching

The phase matching technique [*Jensen et al.*, 2004] does not depend on any assumptions or approximations except the geometrical optics and spherical symmetry of the atmosphere. Meanwhile, there's no tunable parameter in the method such as the position of the back propagation plane or the size of the sliding aperture. The phase matching will be applied along the real satellite orbit, there's no need to assume stationary GPS satellites in the canonical transformation method and circular orbit in the following full spectrum inversion method. In addition, the phase matching function can be given in an explicit form.

If the wave field is given as $u(y) = A(y)\exp(ik\Psi(y)$. Here, y is treated as a generic coordinate such as time t or angle θ. We will replace y in specific coordinates later. As before, $A(y)$ represents the amplitude, $\Psi(y)$ is the eikonal of the wave. $k\Psi(y)$ is the phase of the wave. In the following derivations, we will omit the wave number k in phase functions and use the eikonal. We can always multiply k back in the real computation. Let's consider a signal as a sum of several subsignals:

$$u(y) = \sum_n u_n(y) = \sum_n A_n(y)\exp(ik\Psi_n(y)) \tag{3.32}$$

where n is the number of the interfering subsignals. We introduce a general phase transformation as proposed by [e.g. *Born and Wolf*, 1993; *Jensen et al.*, 2004]:

$$\begin{aligned} F(c) &= \int_0^T u(y)\exp(-ik\Psi_0(c,y)dy \\ &= \sum \int_0^T A_n(y)\exp[ik(\Psi_n(y) - \Psi_0(c,y))]dy \end{aligned} \tag{3.33}$$

where $\Psi_0(c,y)$ is a function of y and can be treated as a GO model. The transformed space is defined by the parameter c. T is the duration of the occultation. The above integral transforms the observation from y domain to c domain. Similar to the generic coordinate y, c can take any value. The technique of phase matching

can also be explained by the stationary phase method [*Born and Wolf*, 1993]. The transformation equation (3.33) is an oscillating integral, the main contribution comes from the vicinity of the stationary phase point. However, if the integrand does not contain any stationary phase point, the result of the integral will be zero. At the stationary point y_s, the frequency of one of the subsignals best matches the frequency of the model $\Psi_0(c, y)$. Thus, we have

$$\frac{d\Psi_n}{dy_s} - \frac{\partial \Psi_0(c, y_s)}{\partial y_s} = 0 \tag{3.34}$$

The frequencies of other signals received at the same point do not match with the model, therefore, no stationary point exists for these subsignals. Those signals have very tiny contributions to the integral. Consequently, we can neglect the sum symbol and the subscript n in (3.33) in the following derivations. Using the stationary phase method, the result of the equation (3.33) is approximately given by [*Jensen et al.*, 2004]:

$$F(c) \approx \sqrt{\frac{2\pi i}{\frac{d^2}{dy^2}[\Psi(y_s) - \Psi_0(c, y_s)]}} A(y_s) \exp[ik(\Psi(y_s) - \Psi_0(c, y_s))]\Big|_{\frac{d}{dy}[\Psi(y_s)-\Psi_0(c,y_s)]=0} \tag{3.35}$$

The phase of $F(c)$ is a function of c. The stationary point y_s can be obtained by taking the derivative of the phase with respect to c. This will lead to a parametric pair:

$$\left[c, \frac{\partial(\Psi - \Psi_0)}{\partial c}\right] \tag{3.36}$$

where

$$\frac{\partial(\Psi - \Psi_0)}{\partial c} = \left(\frac{d\Psi}{dy_s} - \frac{\partial \Psi_0}{\partial y_s}\right)\frac{dy_s}{dc} - \frac{\partial \Psi_0(c, y_s)}{\partial c} = -\frac{\partial \Psi_0(c, y_s)}{\partial c} \tag{3.37}$$

This gives us a functional relationship between the parameter c and the partial derivative of the matching phase.

The phase matching functions $\Psi_0(c, y)$ can be designed to have different forms in order to have desirable properties. If we let $\Psi_0(c, t) = ct$, with $y = t$, this phase matching function gives us a Fourier transformation of the original signal. From equation (3.37), the parametric pair is in the form $[c, -t]$. The expression for c can be related to the impact factor from the Doppler equation:

$$c = \frac{\dot{r}_G}{r_G}\sqrt{r_G^2 - p^2} + \frac{\dot{r}_L}{r_L}\sqrt{r_L^2 - p^2} + p\dot{\theta} \tag{3.38}$$

where r_L and r_G are the distance from the center of the Earth to the satellites. θ is the angle between two satellite vectors. If we assume a circular satellite orbit and a spherical Earth, $\dot{r_G} = \dot{r_L} = 0$. The impact factor is thus proportional to the Doppler

frequency kc. For non-circular orbit, the assumption of the monotonic function $t(c)$ is not necessarily valid. Special procedures are needed to account for the contribution from the non-circular elements.

The form of phase matching function that can deal with the non-circular orbit is given by [e.g. *Jensen et al.*, 2004]:

$$\Psi_0(c,t) = c\theta + \sqrt{r_L^2 - c^2} + \sqrt{r_G^2 - c^2} - c\arccos\frac{c}{r_L} - c\arccos\frac{c}{r_G} \tag{3.39}$$

Geometrically, it represents the sum of two straight distances from LEO and GPS to the tangent points and an arc connecting these two tangent points with angle $\theta - \arccos\frac{c}{r_L} - \arccos\frac{c}{r_G}$ and radius c. The derivative of this function has a similar form to equation (3.38):

$$\omega(c,t) = \dot{\Psi}_0(c,t) = \frac{\dot{r}_G}{r_G}\sqrt{r_G^2 - c^2} + \frac{\dot{r}_L}{r_L}\sqrt{r_L^2 - c^2} + c\dot{\theta} \tag{3.40}$$

The impact factor p can be determined using the relationship $\dot{\Psi}(t_s) = \dot{\Psi}_0(c,t_s)$, which gives $p = c$. This tells us that c actually represents the impact factor. From equation (3.37), the derivative of the transformed phase is

$$\frac{\partial k(\Psi - \Psi_0)}{\partial c} = k\alpha \tag{3.41}$$

Through this phase matching method, we can directly obtain the bending angle as a function of the impact factor. However, phase matching functions given on an arbitrary trajectory are nonlinear, and we cannot perform the integral by the Fast Fourier Transformation (FFT), which makes this method computationally inefficient. The phase matching method is closely related to the full spectrum inversion method. We will introduce their similarity and relationship in the next section.

3.3.4.2 Full spectrum inversion method

The name of the Full spectrum inversion (FSI)[*Jensen et al.*, 2003] is obtained from the fact that a global Fourier Transformation is applied to the entire complex occultation signal (radio hologram) in a circular geometry for deriving the instantaneous frequencies. If we let $\eta = \dot{\Psi}(t) = \frac{d\Psi(t)}{dt}$, the Doppler change ω is equal to $k\dot{\Psi} = k\eta$ and the Fourier Transformation is thus defined as:

$$\tilde{u}(\eta) = \int_0^T u(t)\exp(-ik\eta t)dt \tag{3.42}$$

In the case that the occultation signals consist of several narrow banded subsignals, the FSI method can resolve the frequency variations of each subsignal. The complex signal observed at time t in a circular geometry is denoted as $u(t)$. In a single ray area,

the Doppler frequencies $\omega(t)$ of the rays are the derivative of the phases of the signal. In a multipath zone, however, $\omega(t)$ is contributed to by several interfering signals and can not be simply calculated from the derivative of the phase. Since the Doppler frequency is proportional to the ray impact factor, which is different for different rays, we can assume the same frequency will not occur at different times, i.e., $t(\eta)$ is a single-valued function. This method can be derived and explained by the stationary phase method [*Born and Wolf*, 1993]. To use this method, two requirements must be fulfilled [*Jensen et al.*, 2003], one is the amplitude of each subsignal must vary slowly compared to the phase term, the other is the second order derivative of the phase must be larger than other higher-order derivatives.

Replacing $u(t)$ in equation (3.42) by equation (3.32), yields:

$$\tilde{u}(\eta) = \int_0^T \sum_n A_n(t) \exp(ik\Psi_n(t) - ik\eta t)dt \tag{3.43}$$

The above expression is an oscillating integral, and can be evaluated using the stationary phase method [*Born and Wolf*, 1993]. The phase of the above equation is $k(\Psi_n(t) - \omega t)$. The stationary phase point $t_s(\eta)$ is determined from:

$$\frac{d(\Psi_n(t) - \eta t)}{dt} = \frac{d\Psi_n(t)}{dt} - \eta = 0 \tag{3.44}$$

If the subsignal m fulfills the above equation and dominates the Fourier integral (3.43), the approximate solution of (3.43) is given by the stationary phase method [*Born and Wolf*, 1993; *Jensen et al.*, 2003]:

$$\tilde{u}(\eta) \approx \sqrt{\frac{2\pi i}{k\frac{d^2\Psi_m}{dt^2}(t_s)}} A_m(t_s(\eta)) \exp[ik(\Psi_m(t_s(\eta)) - \eta t_s(\eta))] \tag{3.45}$$

We represent the transformed wave field (3.43) in the form $\tilde{u}(\eta) = A'(\eta)\exp(ik\Psi'(\eta))$. From the solution (3.45), the phase $\Psi'(\eta)$ in the frequency domain is:

$$\Psi'(\eta) = \Psi_m(t_s(\eta)) - \eta t_s(\eta) + \frac{\gamma}{k} \tag{3.46}$$

where $\gamma = \pm\frac{\pi}{4}$ depending on the sign of $\frac{d^2\Psi_m}{dt^2}$. $t_s(\eta)$ is the stationary point from the solution of equation (3.44). However, The last term can be neglected since it vanishes asymptotically. If we take the derivative of the phase of the frequency, we obtain:

$$\begin{aligned} \frac{d\Psi'(\eta)}{d\eta} &= \frac{d}{d\eta}(\Psi_m(t_s(\eta)) - \eta t_s(\eta)) = \left(\frac{\partial\Psi_m(t)}{\partial t}\bigg|_{t=t_s(\eta)} - \eta\right)\frac{\partial t_s}{\partial\eta} - t_s(\eta) \\ &= -t_s(\eta) \qquad \text{with} \qquad \frac{\partial\Psi_m(t)}{\partial t}\bigg|_{t=t_s(\eta)} = \eta \end{aligned} \tag{3.47}$$

If the instantaneous frequency η can occur only once, then the arrival time is treated as a single-valued function. By differentiating the phase with respect to the frequency, the signal arriving time of the frequency is determined and the coordinate of the GPS and LEO satellite can be identified.

The time derivative of the eikonal can be found as [*Jensen et al.*, 2003]:

$$\dot{\Psi} = \eta(p,t) = \frac{\dot{r}_G}{r_G}\sqrt{r_G^2 - p^2} + \frac{\dot{r}_L}{r_L}\sqrt{r_L^2 - p^2} + p\dot{\theta} \tag{3.48}$$

where r_L and r_G are the distance from the center of the Earth to the satellites. θ is the angle between two satellite vectors. If we assume a circular satellite orbit and a spherical Earth, $\dot{r_G} = \dot{r_L} = 0$. Equation (3.48) is simplified to:

$$\dot{\Psi} = \eta = p\dot{\theta} \tag{3.49}$$

$\dot{\theta}$ is a constant. The above equation tells us that frequency is proportional to the ray impact factor. In previous sections, we know that the impact factor can uniquely identify a ray and is different for each ray. This indirectly proves our assumption that any instantaneous frequency can occur only once.

In reality, the orbit of the satellites are not strictly circular, and the orbital plane of the GPS and LEO does not coincide. Meanwhile, due to the oblateness of the Earth, the center of the refractivity is generally not the center of the Earth. As a result, we cannot assume the radial velocities of the satellites are zero and θ is no longer a simple linear function of time when two satellites are not in the same plane. In such situations, the monotonic relationship between the time and frequency might be broken since the radial accelerations and higher-order derivatives of θ make the FSI result too noisy.

In the above discussion, for the circular orbital configuration, the FSI method can be treated as a special case phase of matching method with $\Psi(\eta, t) = \eta t$, $c = \eta$, and $y = t$. It is indeed a Fourier method as discussed in the phase matching section. The phase matching method can account for the non-circular orbit, but function (3.39) is not linear and can not use the FFT method. The phase function (3.39) can be revised to approximate the original function but with the linear form. It can be written as a function of θ:

$$\Psi_{FSI}(c, c_0, \theta) = c\theta + \sqrt{r_L^2 - c_0^2} + \sqrt{r_G^2 - c_0^2} - c_0 \arccos\frac{c_0}{r_L} - c_0 \arccos\frac{c_0}{r_G} \tag{3.50}$$

Thus, from equation (3.37), we obtain the following parametric pair:

$$\left[c, \frac{\partial(\Psi - \Psi_{FSI})}{\partial c}\right] = [c, \theta_s] \tag{3.51}$$

The impact factor p can be found from the Doppler equation $\omega(\theta_s) = \omega_{FSI}(c, c_0, \theta_s)$, i.e.,

$$\frac{\dot{r}_G}{r_G}\sqrt{r_G^2 - p^2} + \frac{\dot{r}_L}{r_L}\sqrt{r_L^2 - p^2} + p\dot{\theta} = \frac{\dot{r}_G}{r_G}\sqrt{r_G^2 - c_0^2} + \frac{\dot{r}_L}{r_L}\sqrt{r_L^2 - c_0^2} + c\dot{\theta} \tag{3.52}$$

p is not equal to c comparing the phase matching method. With the proper choice of c_0, the impact of the radial velocity can be reduced. Finally, the bending angle can be calculated as:

$$\alpha = \arcsin(\frac{p}{r_G}) + \arcsin(\frac{p}{r_L}) + \theta_s - \pi \tag{3.53}$$

In the real computation, the radio occultation observation needs to be up-sampled and interpolated with respect to θ coordinates.

3.3.5 The relationship among CT, FSI and phase matching

In this section, the relationship among CT, FSI, and phase matching methods will be discussed. These methods are theoretically related and some of them can be mutually explained. We already understand in the above section that the FSI method is a phase matching method with a linear phase matching function, which enables the use of the FFT for better computational efficiency. All these methods transform the signal to another form of representation which is capable of disentangling the interfering rays in the multipath region, and they all use the derivative of the transformed phase to compute the bending angle and impact factor. On the other hand, they all have some restrictions, such as CT needs a time-consuming pre-processing step to back-propagate the signal to a vertical line, FSI works best in a circular orbital configuration, and the phase matching method can work in arbitrary observation trajectories but FFT cannot be applied. In the following, we interpret these methods from alternative angles.

3.3.5.1 CT and phase matching principle

We already know that the stationary phase method can be applied both in the derivation of the canonical transformed method and phase matching method. We will show here that the canonical transform method can also be derived by the phase matching principle.

According to equation (3.33), the phase matching function in the frequency domain can be defined as:

$$\tilde{F}(p) = \int \tilde{u}(\eta) e^{-ik\tilde{\Psi}_0(p,\eta)} d\eta \tag{3.54}$$

with the phase matching function for frequency representation:

$$\tilde{\Psi}_0(p,\eta) = -\int y(p,\eta) d\eta \tag{3.55}$$

It has a similar form to the generating function $S(p,\eta)$ in equation (3.26), but with opposite sign. Consequently, we can use the expression for the generating function:

$$\tilde{\Psi}_0(p,\eta) = -p \arcsin\eta + x\sqrt{1-\eta^2} \tag{3.56}$$

From equation (3.37), the derivative of the transformed phase with respect to p is:

$$\frac{\partial(\Psi - \Psi_0)}{\partial p} = -\frac{\partial \Psi_0(p, \eta_s)}{\partial p} = \arcsin \eta = \xi \tag{3.57}$$

For a given p, we can compute the ray arriving angle from the above equation, this is the same equation (3.27) from the CT method.

Substituting equation (3.56) into equation (3.54), yields:

$$\tilde{F}(p) = \int \tilde{u}(\eta) e^{-p \arcsin \eta + x\sqrt{1-\eta^2}} d\eta \tag{3.58}$$

It is similar to the FIO equation (3.21) in the CT method:

$$u_x(p) == \frac{k}{2\pi} \int a(p, \eta) e^{ik(p \arcsin \eta - x\sqrt{1-\eta^2})} \tilde{u}_x(\eta) d\eta$$

Only the amplitude functions between these two methods are a little bit different. From the above derivation, we clearly see that the CT method can also be interpreted by the phase matching principle in the frequency domain.

3.3.5.2 CT and FSI method

The FSI method is a special form of the phase matching method. As shown above, the CT method can be formulated by the phase matching principle used by the FSI method. Reversely, the FSI method can be re-formulated by the CT method, which means FSI is also a CT method but with different coordinate representations. The FSI method is along a satellite orbit which is close to a circle, while the CT method is applied to the fields on a back-propagated straight line. The transformed field of the two methods are different. For the FSI method, the transformed field is a function of Doppler frequency which is approximately proportional to the ray impact factor. While for the CT method, it is a function of the impact factor directly. Numerically, the FSI method is relatively simple, easy to implement, and computationally efficient, but it requires a small orbital eccentricity.

3.3.5.3 The comparison of radio-holographic methods

In Table 3.1 , the characteristics of different radio holographic methods are compared and summarized.

Radio holographic methods	Tunable parameter	Transmitter position	Observ. trajectory	FFT	Derived by stationary phase method	Derived by FIO	Phase matching	Remove Diffraction	CPU power
Back Propagation	Position of BP plane	Stationary	Vertical line	Yes	N/A	N/A	N/A	Yes	High
Canonical Transform	No	Stationary	Vertical line	Yes	Yes	Yes	Frequency domain	Yes	High
Phase Matching method	No	Non-stationary	Arbitrary	No	Yes	N/A	Coordinate or time domain	No	Low
Full Spectrum Inversion	No	Non-stationary	Circular orbit	Yes	Yes	Yes	Yes	No	Low
Fourier Integral type 2	No	Non-stationary	Arbitrary	Yes	Yes	Yes	Yes	No	Moderate
Radio Optic	Size of sliding aperture	Non-stationary	Arbitrary trajectory	Yes	N/A	N/A	N/A	No	Low

Table 3.1: Comparisons of radio holographic methods

CHAPTER 4

RESULTS FROM THE GPS OCCULTATION AND THEIR VALIDATIONS

4.1 Introduction

In order to better understand the GPS occultation retrieval process and better utilize the high level products, we have been continuously improving our own retrieval software to obtain accurate atmospheric parameters. In this chapter, we discuss the characteristics of GPS occultation and its potential to improve the pressure fields. The retrieved occultation profiles are validated against the numerical weather prediction products, mainly from ECMWF and NCEP.

4.2 The characteristics of GPS occultation

4.2.1 Spatial coverage

The number of occultation events for a particular satellite mainly depends on its orbital configurations (e.g. altitude, inclination). The GPS antenna for occultation is pointing in the direction along the satellite track. Satellites can carry both fore-looking and aft-looking antennae (e.g. SAC-C) for occultation purposes. The events are usually called rising and setting events, respectively, and last about one minute. In practice, however, rising events require accurately predicting and tracking the signals from the very beginning without the assistance of any prior tracking knowledge, which is very hard using the current hardware architecture – closed Phase Lock Loop (PLL). An open loop algorithm is required and is currently under active development. Therefore, almost all occultation events currently available are setting occultation events.

The GPS occultation system is an all-weather system, and it is insensitive to cloud and rain. The basic measurement is the time-delay of the signal when passing through the Earth's atmosphere. The occultation events have nearly global coverage. Typically in one day, there are roughly 250 occultation events. After overcoming

the difficulty of tracking rising occultations, this number can be doubled. Figure 4.1 shows an example of the spatial coverage for a typical day. The sounding points are nearly uniformly distributed.

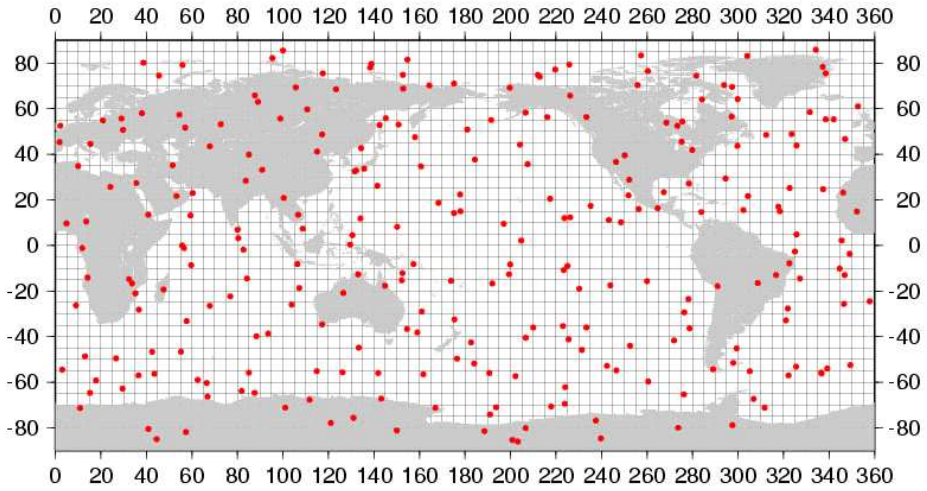
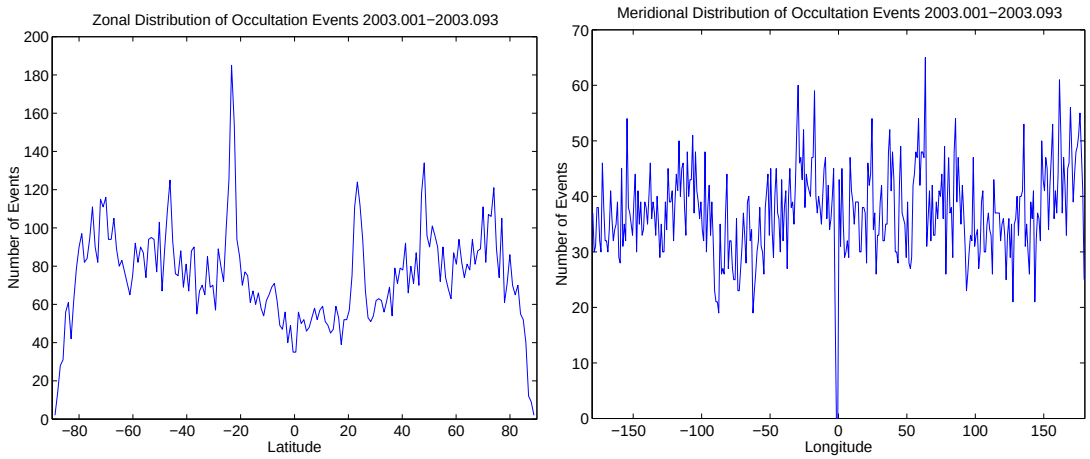


Figure 4.1: Locations of GPS occultation events for one day.

In reality, the occultation events are not strictly uniformly distributed. Due to the orbital configuration of the GPS satellites, there are more occultation events in the middle latitude area than in the tropical and polar regions (Please note that the zonal area becomes smaller when moving from tropics to polar regions). Meridional distribution is more uniform. The distribution of 3 months of CHAMP occultation events is shown in Figure 4.2. The events are counted using a 1 degree bin size along the longitude and latitude, respectively.

In the data sparse regions, such as the Southern Oceans and Antarctica, the shortage of regular types of observations limits the model performance. As a new type of observation, the GPS occultation, theoretically has the potential to remedy the data deficiency by increasing the spatial coverage. Figure 4.3 shows the geographic coverage of three different measurement techniques (CHAMP occultation Jan.-Mar., 2003, automatic weather stations, and radiosondes). It is obvious that these occultations can provide more geographical coverage over Antarctica compared to the sparsely located radiosonde stations (15 stations for the entire Antarctic region), while the other two data types provide dense time series but at fixed locations.



(a) Zonal (b) Meridional

Figure 4.2: Zonal and meridional distributions of 3 months CHAMP occultation events.

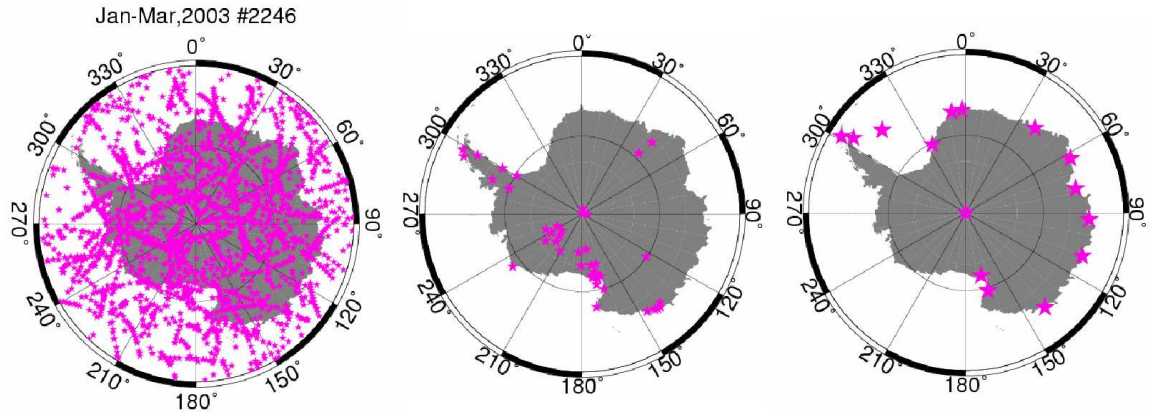


Figure 4.3: Coverage maps for CHAMP occultation Jan.-Mar., 2003(left), automatic weather stations (middle) and radiosonde (right) in Antarctica.

4.3 CHAMP profile retrieval

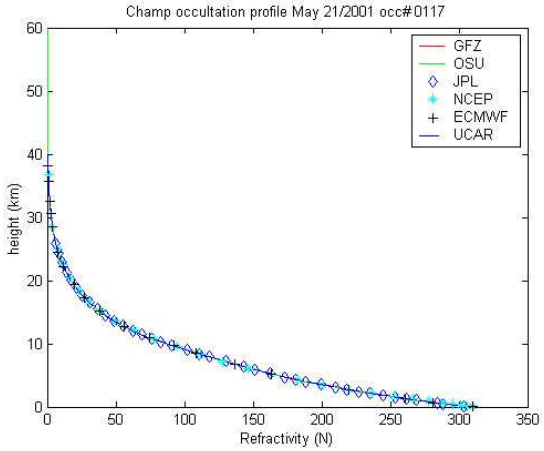
We chose a CHAMP event near a radiosonde station over Antarctica. Figure 4.4 shows the computation results based on the algorithms described in previous sections. The profile (18.4°W, 70.1°S) was recorded on May 21, 2001, 12:54 UTC

and the data are obtained from the CHAMP data center at GFZ. Level 2 product (atmospheric excess phase) is used to compute the vertical atmospheric profiles, which are compared to profiles from other data centers (GFZ level 3 product, JPL level 2 product, UCAR COSMIC Data Analysis and Archive Center (CDAAC) level 2 product) and corresponding ECMWF and NCEP data at the same location. From the comparison of the temperatures (Figure 4.4c), we see that the occultation results agree well among the Ohio State University (OSU) solution, UCAR and GFZ solutions below 30 km. The large disagreement of the temperature profile above 35 km between OSU and GFZ solutions is primarily due to the different methods adopted for the boundary condition of the upper atmosphere. The OSU solution chooses a statistical optimization method above a certain altitude, usually around 35-40 km. At those altitudes, noises dominate bending angle measurements. The GFZ retrieval software applies another approach for the optimization. This is the reason for the deviations observed above 30 km. JPL results agree with OSU, GFZ, and UCAR for 5-25 km, and they fit well with the NCEP and ECMWF data in the lower troposphere. All the results from other processing centers also show good agreement with the NCEP and ECMWF analysis results except in the lower 5 km. We note that the GPS occultation measurement has a much higher vertical resolution than the weather analysis products have.

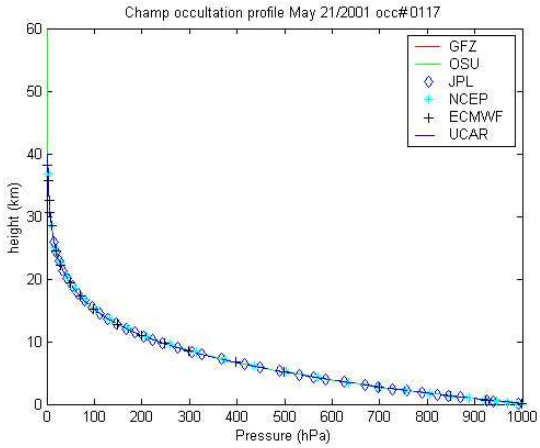
GPS occultation can also be validated by other measurement techniques, like radiosonde. To demonstrate the potential of the GPS radio occultation technique for precise and high vertical-resolution temperature profiling above Antarctica, the CHAMP occultation profile (Figure 4.4) is compared with radiosonde data. The radiosonde was launched at the German Neumayer station (8.2°W, 70.4°S) at 9 UTC on May 21, 2001. The data were recorded between 9 and 11 UTC. The agreement between both profiles is very good above the Earth's surface (temperature here are about -35°C) up to about 10 hPa pressure level (30 km height). The precise resolution of the tropopause height and temperature (Figure 4.5) is remarkable. It indicates that the occultation technique has comparable accuracy with radiosonde data. Statistical validations using larger data sets from CHAMP and SAC-C missions, as well as radiosonde and meteorological analysis data are performed in a later section to assess the potential use of GPS occultation data for improving the pressure modeling in Antarctica.

4.4 Signal penetration

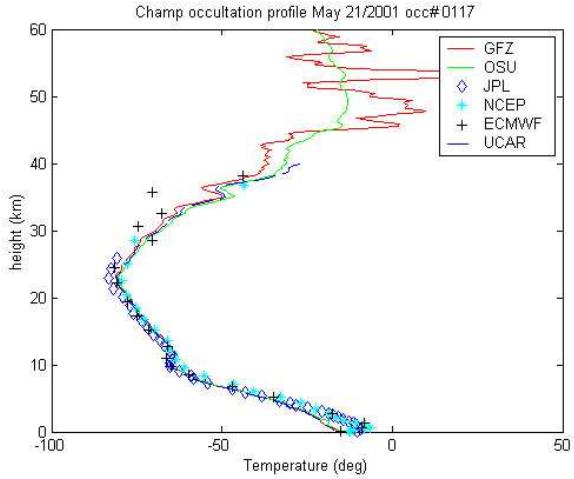
As discussed in previous chapters, there are many factors that affect the penetration depth of occultation profiles, such as hardware sensitivity, tracking algorithms, environmental diffractions, and multipath (due to water vapor). To accurately derive the surface pressure, occultation signals need to penetrate down to the Earth's surface. Due to the complexity of water vapor and signal tracking problems in the



(a) Refractivity profile



(b) Pressure profile



(c) Temperature profile

Figure 4.4: A CHAMP profile over Antarctica.

Planetary Boundary Layer (PBL), signals sometimes cannot penetrate down to the surface. In tropical regions, the water vapor is abundant in the atmosphere; while in the cold polar region, the air hardly contains moisture. The moisture content could induce a very strong refractivity gradient, and the tracked signals are noisier than in the dry air. In severe situations, the signal loss-of-lock could happen. We chose two occultation events, one is at a tropical region, the other is at a polar region. Figure 4.6 shows that signals tracked in the tropical region have lower signal-to-noise ratios, and they become much noisier when approaching the ground. The derived excessive

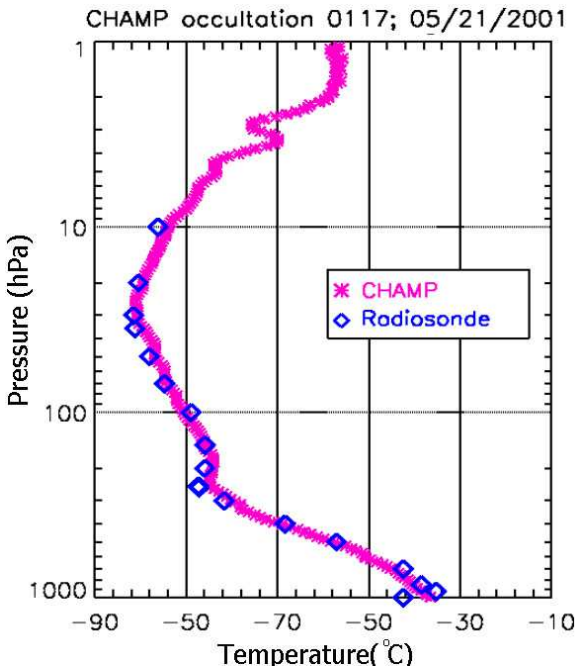
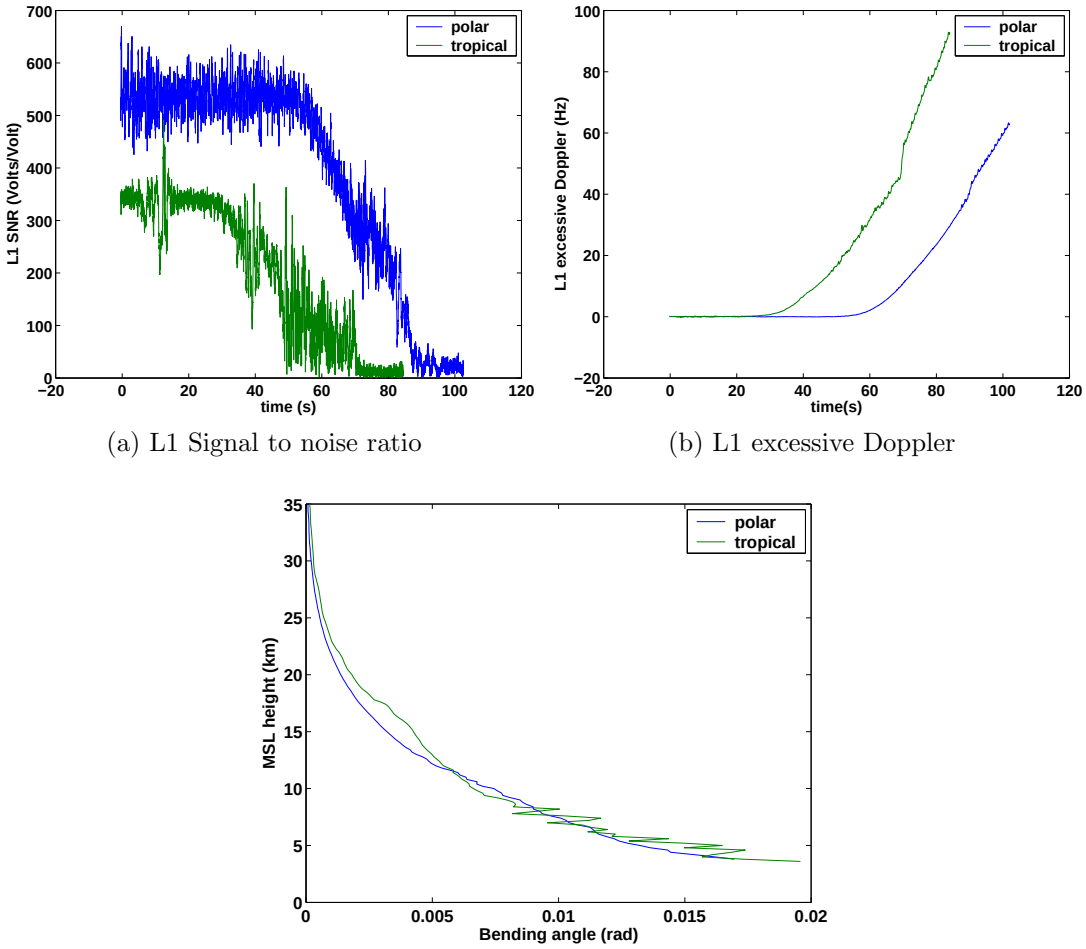


Figure 4.5: Comparison of CHAMP vertical dry temperature profile with radiosonde data from the German Neumayer station on May 21, 2001.

Doppler is relatively smooth in polar regions. Consequently, the derived bending angle is also relatively smooth. The results in such region are more reliable and easier to interpolate and extrapolate to the ground.

We use 3 months of data to investigate the signal penetration to the ground. Figure 4.7 illustrates the distribution of CHAMP occultation penetration over three different regions. In the tropical region (defined here as 30°N-30°S), only about 10% of signals penetrate down to 1 km above the MSL (Figure 4.7b). This is caused by the warm and moist environment in the tropical region. The rapidly changing moist content in the lower atmosphere also affects signal tracking. CHAMP occultation performs much better in the Arctic region (Figure 4.7a), nearly 80% of the profiles reach within 1 km above the MSL. The deeper penetration results because (1) most of the Arctic region is over the ocean; and (2) the Arctic region is cold and relatively dry. Figure 4.7c shows the results over Antarctica, where nearly 50% of profiles reach 1 km above MSL. It is noted that Antarctica is a highly elevated region. After removing the topographic effect (i.e., referencing the occultation profiles to the ECMWF topography), Figure 4.8 shows that the signal penetration in Antarctica is similar to that in the Arctic region (80% of signals penetrate to within 1 km above the surface).



(a) L1 Signal to noise ratio

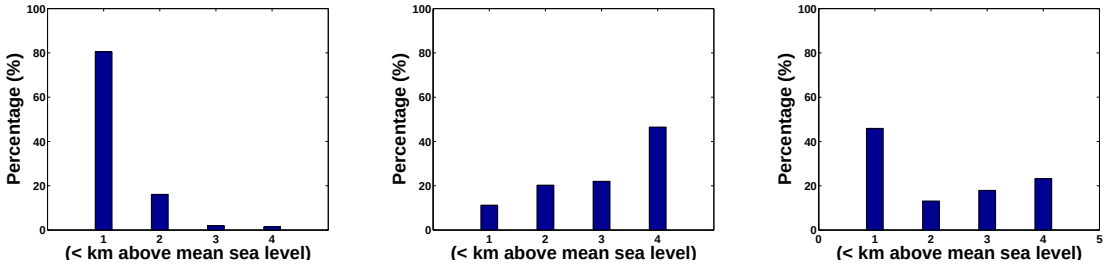
(b) L1 excessive Doppler

(c) Ionosphere-free bending angle

Figure 4.6: The comparison of polar and tropical GPS occultation events.

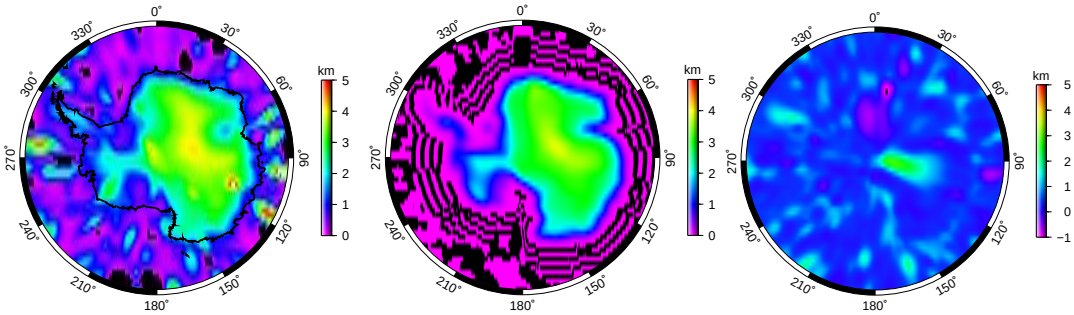
4.5 The comparison of CHAMP and SAC-C pressure profiles

During the past several years, there has been great developments in the GPS occultation retrieval algorithms, especially for profiles in lower tropospheric multipath regions. Data processing centers, e.g., JPL, UCAR, and GFZ, are continuously providing low and high level GPS occultation products. Differences still exist in the retrieval algorithms among data centers. Data from different centers should be accurate, consistent and comparable. Testing the consistency of these products, especially the pressure field, will guide us to properly utilize the data.



(a) Arctic (60°N-90°N) (b) Tropical (30°S-30°N) (c) Antarctic (60°S-90°S)

Figure 4.7: Histograms of CHAMP occultation penetration depths (referencing to the MSL) over the Arctic (80% signal penetration), tropical (10%), and Antarctic (50%) regions. CHAMP occultation data are from Jan.-Mar., 2003.



(a) Above MSL (b) ECMWF topography (c) Above the Earth's surface

Figure 4.8: The lowest penetration altitude above different penetration surfaces.

The co-existence of CHAMP and SAC-C provides a unique opportunity to examine the inter-satellite consistency. In this particular case study, we compare the occultation derived pressure profiles at a specific altitude (here, it is 10 km). Occultation pressure profiles from JPL and UCAR based on identical datasets are analyzed with $10° \times 10°$ block averages on the entire Earth. We will not be able to match positions between CHAMP and SAC-C profiles. The comparison between CHAMP and SAC-C is done indirectly by subtracting occultation derived pressure from each satellite from an identical NWP analysis model. The inter-satellite consistency is evaluated by comparing the two differences (CHAMP vs. model and SAC-C vs. model). Since CHAMP and SAC-C have different orbit configurations, this result can show us whether the occultation result is mission independent in a relatively short time span

(e.g., one month). How large a discrepancy of occultation results from independent data centers can be expected is another question to answer.

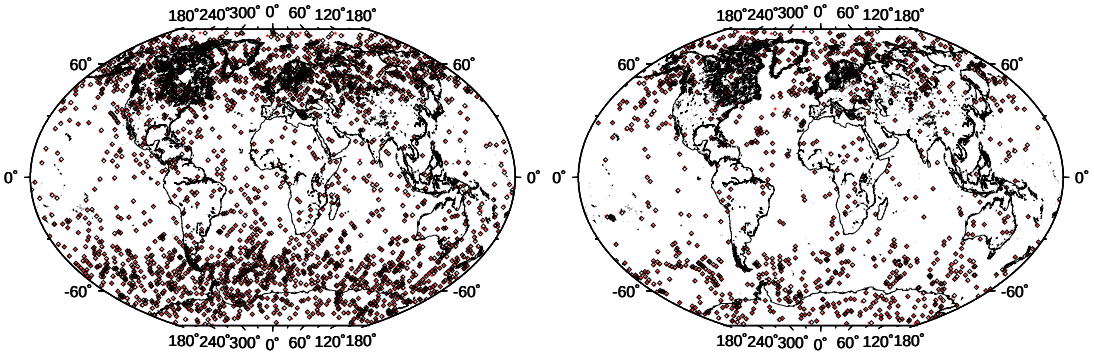
4.5.1 Datasets

The period chosen in this study is Sept. 1-30, 2001. The reason for choosing this period is simply that CHAMP and SAC-C have the same tracking software, but CHAMP has a later version of a scheduler while SAC-C has a later version of the master control program (JPL GENESIS website). The level 2 data were downloaded separately from the UCAR COSMIC website (http://www.cosmic.ucar.edu) and the GENESIS ftp site (ftp://sayatnova.jpl.nasa.gov). Only profiles that both centers have are used in this study. There are a total of 1,221 CHAMP profiles and 2,362 SAC-C profiles for the analysis (Figure 4.9). ECMWF and NCEP profiles are obtained by interpolating along the occultation path. For pressure measurements, the linear log interpolation was used along the altitude. Figure 4.9 shows the occultation coverage during the whole analysis period. Since SAC-C has a relatively higher orbit altitude (702 km) than CHAMP (initial orbit 454 km), it has more GPS occultation events than CHAMP. Due to the orbit configuration, mid-latitude and polar regions have more occultation events than the equatorial region.

CHAMP and SAC-C pressure profiles along with ECMWF and NCEP pressure profiles are interpolated between 1 km and 30 km altitude above mean sea level in 1 km step. The depth of the CHAMP profile penetration affects the accuracy of the derived surface pressure. Since the pressure usually increases exponentially with a decrease in altitude, the extrapolation error also increases exponentially. To avoid the effects of the lower troposphere multipath and the water vapor ambiguity to the analysis result, we compare the pressures at 10 km altitude.

4.5.2 Comparison results

The differences between the CHAMP occultation derived pressure and ECMWF pressure at 10 km altitude are shown in Figure 4.10. The differences are averaged using the $10^\circ \times 10^\circ$ block mean and interpolated over the globe. We assume that the model analysis results can remove most of the time and location dependent pressure values from occultation profiles. Panel (a) shows the results from two centers. JPL CHAMP-ECMWF difference has several negative anomalies ($<$ -3 hPa) around the equatorial region, while UCAR CHAMP-ECMWF difference has positive anomalies ($>$ 4 hPa) around the southern mid-latitudes. The differences in most of the region in these two figures are within $\pm$1 hPa. The two figures of CHAMP-NCEP differences from UCAR and JPL also look slightly different from each other, especially at the southern polar regions. If comparing the two figures from the same organizations in Figure 4.10, we can find some common anomalies. Their magnitudes are slightly different due to the difference between NWP models. In the Southern Ocean and



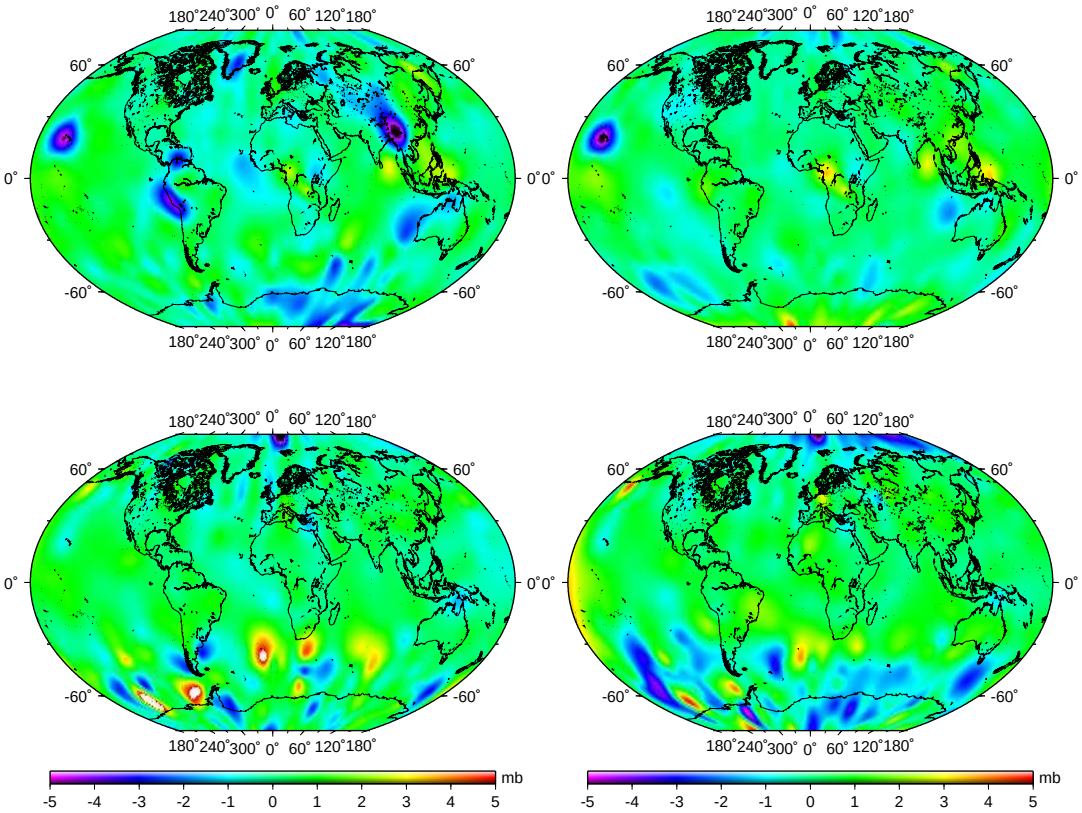
(a) SAC-C (b) CHAMP

Figure 4.9: Locations of the GPS occultation profiles (2,362 for SAC-C and 1,221 for CHAMP).

Antarctica, JPL's results show that CHAMP pressure is lower than ECMWF and higher than NCEP. UCAR's results tell the opposite, and the CHAMP pressure is lower than NCEP in such regions. These disagreements reflect the possible influence of the retrieval algorithms on the occultation derived pressure, and on the large disagreement between CHAMP and model in Southern Ocean and Antarctica.

The differences between SAC-C data and models are more prominent than that of CHAMP (Figure 4.11). SAC-C pressure profiles from JPL have positive equatorial differences larger than 5 hPa, which can be seen in both ECMWF and NCEP comparisons (Top two figures in Figure 4.11). This implies that these anomalies probably originate from SAC-C profiles. As for UCAR's retrieval results, several negative anomalies occur near the equatorial and mid-latitude regions in both ECMWF and NCEP figures (Bottom two figures in Figure 4.11). Among these anomalies, the occultation pressure is lower than the model pressure. Here once again we see disagreements between the JPL's and UCAR's results. In some areas of Antarctica, JPL's SAC-C pressure is lower than ECMWF and higher than NCEP. UCAR's SAC-C pressure has a smaller value compared to NCEP in the Southern Ocean and Antarctica, where the pattern looks similar to the UCAR's CHAMP and NCEP comparison. This probably suggests that NCEP performs worse than ECMWF at the southern polar region.

The occultation pressure results are inconsistent between different processing centers. Meanwhile, different satellites do not produce a consistent error pattern. For JPL's solution, CHAMP pressure seems lower than SAC-C pressure. UCAR's solution does not have a clear pattern between CHAMP and SAC-C, and the signs and locations of the anomalies are also different.

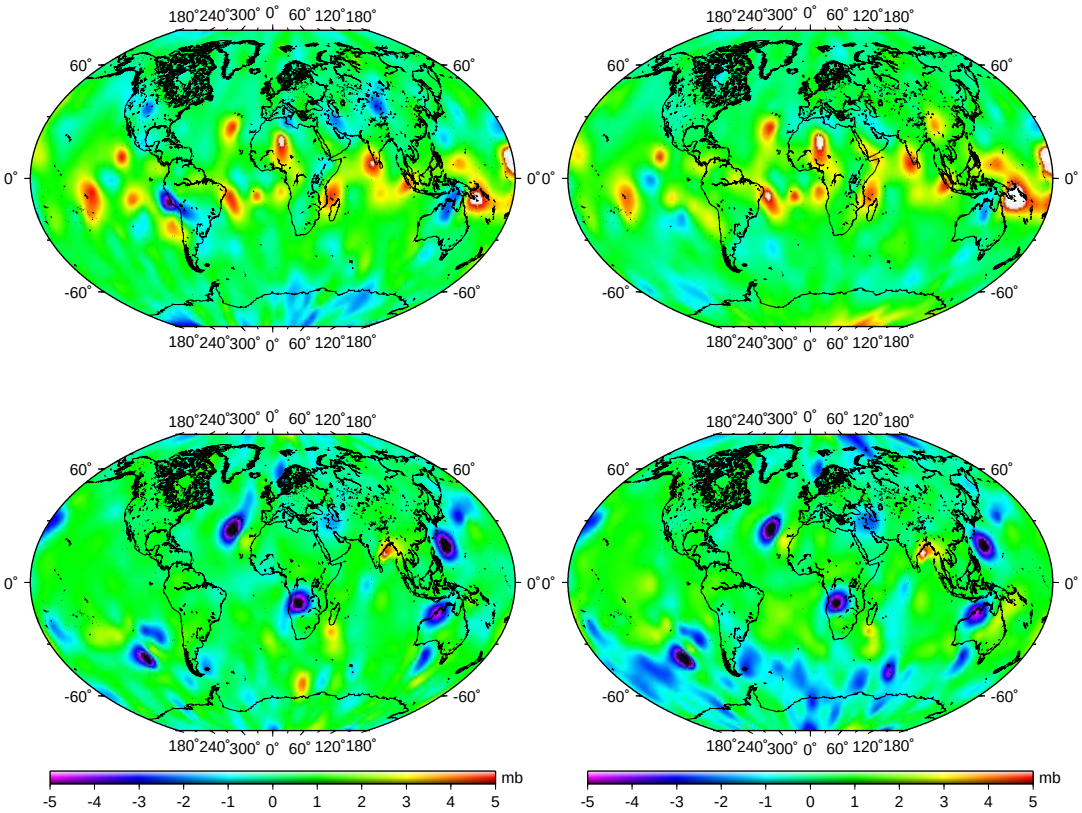


(a) CHAMP - ECMWF (b) CHAMP - NCEP

Figure 4.10: Pressure differences between CHAMP occultation profile and analysis models at the 10 km altitude. (Derived by $10^\circ \times 10^\circ$ block mean, JPL(top), UCAR(bottom)).

The above results show inconsistency between different centers and between different satellites. At 10 km altitude, there is no multipath and no water vapor ambiguity problem, but the pressure difference sometimes exceed 5 hPa. In UCAR's results, large disagreements with NCEP are around the Antarctic region. JPL's SAC-C pressure results show big equatorial difference with both models. The reason for these differences is at present unclear. All the profiles used in this study have passed quality control.

When comparing each single profile from different centers, we find that they usually do not agree with each other perfectly. Pressure has a relatively larger range compared to temperature or water vapor (one order of magnitude larger). It also increases or decreases exponentially with the change of height. The cause of inconsistency between different centers might be due to different retrieval algorithms, initialization



(a) SAC-C - ECMWF (b) SAC-C - NCEP

Figure 4.11: Pressure differences between SAC-C occultation profile and analysis models at the 10 km altitude. (Derived by $10° \times 10°$ block mean, JPL(top), UCAR(bottom)).

and optimization methods, and different approximations adopted in the processing procedures. Assessing these procedures and approximations and their sensitivities to the pressure retrieval can help us to better understand how these differences could occur. Especially, the steps that are related to the definitions or approximations of geopotential height need to be carefully validated.

Current ECMWF analysis is already working very well compared to other models. If one wants to improve the models using occultation measurements, errors and inconsistencies must be identified. A stringent quality control procedure is also required.

4.6 Comparison of GPS derived pressure with ECMWF, NCEP and radiosonde data

4.6.1 GPS occultation dataset

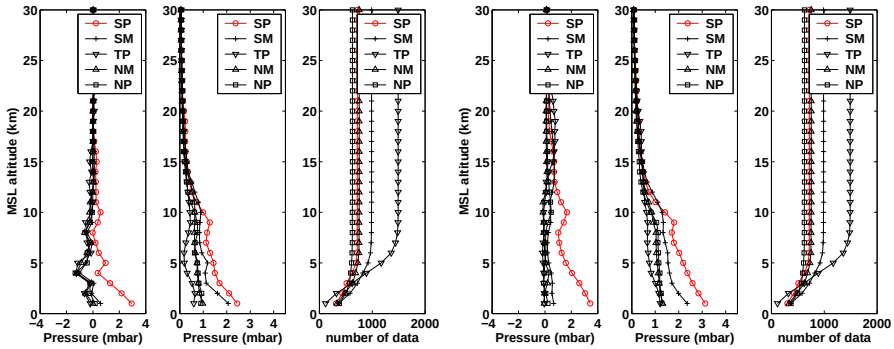
The period chosen in this study is April, 2003 (or day 091-120, 2003). The data are obtained from the COSMIC data center. ECMWF and NCEP profiles are obtained by interpolating along the occultation path. If we consider CHAMP and ECMWF only, there are a total of 4,574 matched profiles, while for radiosonde data, only 1,279 matched profiles exist.

4.6.2 Method

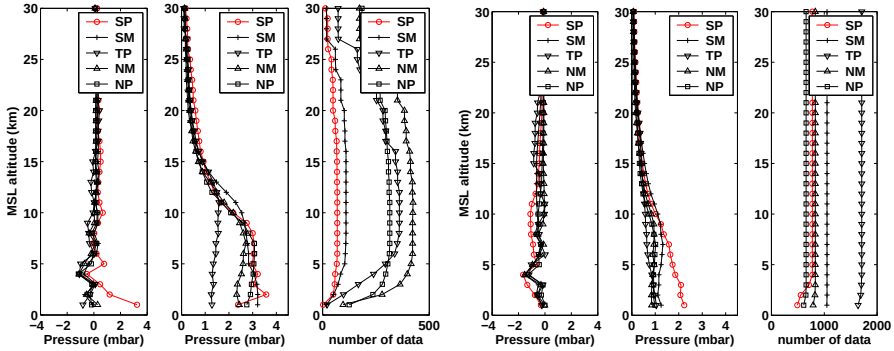
Before the comparison, the CHAMP, ECMWF, NCEP, and radiosonde pressure profiles are interpolated between 1 km and 30 km altitude above MSL in 1 km step. No extrapolation is performed if the data did not reach a certain altitude, e.g., 1 km or 2 km. The comparison is done to the levels where each two types of data exist. Since usually models are expected to perform differently in different regions, we divide the whole globe into 5 regions according to the latitudes: Southern Polar (SP, 60°S-90°S), Southern Mid-latitude (SM, 30°S-60°S), Tropical (TP, 30°S-30°N), Northern Mid-latitude (NM, 30°N-60°N), and Northern Polar (NP, 60°N-90°N) regions. Since the extrapolation error increases exponentially with a decrease in altitude, the depth of the CHAMP profile penetration affects the accuracy of the derived surface pressure. The mean and standard deviation (STD) are calculated for each pair of data in the analysis period. In order to compare with radiosonde data, many profiles in the other datasets have to be discarded simply because radiosonde measurements only exist over land area. There are only a few match-ups in the southern hemisphere (Figure 4.12, right panels). Without losing generality, we compute and compare the statistics of 4,574 and 1,279 matched CHAMP and ECMWF profiles (reduced to 1,279 in order to match radiosonde profiles). Since the statistics behave similarly, we ignore the different number of observations when we discuss our results.

4.6.3 Comparison results

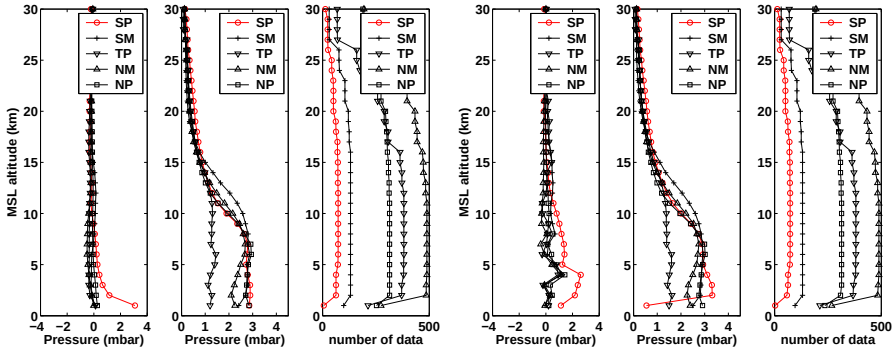
The differences among 4 types of pressure profiles are shown in Figure 4.12. We compare CHAMP with ECMWF, NCEP and radiosonde pressure profiles in the first sets of panels in Figure 4.12. From the left panel of (a), (b), and (c), it is obvious that the mean CHAMP pressure at the SP region is lower than that of any of the other 4 regions, or CHAMP is biased negatively versus the other models or data. It is nearly 3 hPa at 1 km altitude for the Antarctic region (SP). The biases for other regions in (a) and (b) are within 1 hPa at 1 km altitude. Larger STD differences for the SM and SP regions at 1 km altitude are shown in the middle figures in (a)



(a) $P_{ECMWF} - P_{CHAMP}$ (b) $P_{NCEP} - P_{CHAMP}$



(c) $P_{SON} - P_{CHAMP}$ (d) $P_{ECMWF} - P_{NCEP}$



(e) $P_{ECMWF} - P_{SON}$ (f) $P_{NCEP} - P_{SON}$

Figure 4.12: Inter-comparisons (pressure difference) of CHAMP derived pressure profile (P_{CHAMP}), ECMWF pressure profile (P_{ECMWF}), NCEP pressure profile (P_{NCEP}) and radiosonde pressure profile (P_{SON}). In the figures: SP–southern polar region (60°S-90°S), SM–southern mid-latitude region (30°S-60°S), TP–tropical region (30°S-30°N), NM–northern mid-latitude region (30°N-60°N), NP–northern polar region (60°N-90°N). For each figure: left panel–mean of pressure difference, middle panel–standard deviation (STD) of pressure difference, right panel–number of values in each level used for the comparisons.

and (b). STD differences at 1 km altitude in all other regions are generally around 1 hPa. It is surprising to find that the tropical region (TP) has the smallest STD. In the comparison of the ECMWF pressure with the CHAMP pressure, the agreement in the SP region is the worst both in bias and STD.

In (b) of Figure 4.12, the large bias between NCEP and CHAMP pressure profiles also occurs around the tropopause (SP). The SP area still has the largest bias at 1 km altitude. It is expected to be even larger at the Earth's surface. The STD figure shows a similar pattern to that of ECMWF. The tropical region (TP) still has the smallest STD. The similarity of (a) and (b) directly causes the outputs of ECMWF and NCEP to agree relatively well as shown in (d). The bias is within 1 hPa and the STD difference for all regions are around 1 hPa except for the SP region. The largest bias occurs at around 4 km altitude. Although the STD is relatively small compared to other figures, SP is still the region which has the largest discrepancy between these two models.

As an independent measurement technique, radiosonde is a valuable data source in this study. The remaining figures (c, e, and f) in Figure 4.12 compare the radiosonde pressure with the pressure from CHAMP, ECMWF, and NCEP. The figure in the left panel of (c) could probably confirm that the biases between CHAMP and ECMWF and between CHAMP and NCEP come from the CHAMP pressure profiles. The STD figure in (c) also shows big value at the bottom, and the TP region has the smallest value. We should point out that the match between the occultation and the radiosonde profile cannot be exactly both in location and time, i.e., the match is only defined within a certain period of time (e.g. ± 1 hour) and within a certain distance (e.g. 200 km). The comparisons between ECMWF/NCEP and radiosonde have the same problem. The variation of pressure is usually larger in the mid-latitude and in polar regions than in the tropical region (see Chapter 5, Figure 5.2). Therefore, the mis-match of time and/or position of two profiles will increase the STD of all regions with the SP region as the highest. This is confirmed by the middle panel of (c), (e), and (f).

Figure (e) and (f) in Figure 4.12 show that there are large mean differences between the model pressure and radiosonde pressure in the SP region. This is reasonable because that the lack of data in the southern polar region causes poor model performance. The differences are larger near the surface. One should note that the number of radiosonde stations in Antarctica is very small (see right panel of (c), (e), and (f)), the statistics for such a region should be carefully interpreted.

Through the analysis of Figure 4.12, we find that the pressure bias is large at or near the Earth surface, especially in the SP region. CHAMP derived pressure has smaller value in each level compared to other data in the SP region. From the STD of CHAMP and radiosonde difference, we can see that the two measurement techniques do not agree very well. One reason could be that the radiosonde has lower vertical resolution than the occultation does; the other possibility is that they do not

exactly match both in space and time. The biggest advantage of GPS occultation is its spatial coverage. Once the biases are removed and the errors are detected, it can be used to improve the model performance, especially in data-sparse regions. We will show in the next section that we could use techniques such as one-dimensional variational (1DVar) method to combine occultation and model values in an optimal way to improve the accuracy.

4.7 1DVar method

The variational approach is a method to obtain a statistically optimal atmospheric state by combining a background model (*a priori*) with observations. We have seen that GPS occultation has many attractive characteristics. It has high vertical resolution and accuracy. The sounding points have global coverage and are almost unaffected by precipitations and clouds. Unlike the radiosonde, its instrument requires no calibration, no instrument drift and no bias exist for different satellites. However, GPS occultation observation represents a new type of observation, which is not a traditional meteorological measurement and cannot be directly used. The vertical profile is not at a single position. Instead, it is along a slant path of about 200-300 km. Even a single point in a vertical profile can not be simply viewed as a local measurement. It is shown that converting the path delay to meteorological quantities is subject to various assumptions and errors (e.g., spherical geometry, choice of optimization methods, height of initialization, etc). The results are sensitive to various algorithms.

To avoid the above mentioned problems, one can choose to assimilate the quantities obtained in the early steps of the profile retrieval. The possible choices include (1) excessive phases and/or amplitudes; (2) bending angles; (3) refractivity; and (4) retrieved pressure, temperature and humidity profiles. The model assumptions and approximations have almost no effect on the raw data (1), but this method requires a complete modeling of wave propagations and ionospheric effects, as well as the satellite orbital information. It is computationally expensive and impractical. Choice (2) does not require the orbital information any more, but needs a 3-D ray tracing operator to connect the bending angle and numerical weather analysis quantities [*Zou et al.*, 1999]. It is also computational demanding. Choice (3) only requires an upper boundary initialization, and the observation operator is relatively easy, though there are errors in this choice. The implementation is relatively inexpensive. Choice (4), obviously, is not a good candidate, because it includes all the errors in the algorithms.

Assimilating the occultation profiles to real numerical weather models requires the development of a 3-dimensional or 4-dimensional variational system, which is out of the scope of this research. Here, we want to present a simulation result to combine a radio occultation profile with a profile from an analysis model in the framework of

1DVar to solve the "water vapor ambiguity" and assess the impact of occultation on the analyzed surface pressure. We use the refractivity as observation in our simulation.

The penalty function for 1DVar is set up in the form [*Eyre*, 1994; *Healy and Eyre*, 2000]:

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}(y_0 - H(x))^T (E + F)^{-1}(y_0 - H(x)) \tag{4.1}$$

The mathematical task here is to minimize the penalty function. x in this case is a vector including a profile of pressure, temperature and water vapor (i.e. the most probable atmospheric state). x_b is the state vector of the atmospheric background (*a priori*, derived from the background information). B is the covariance matrix of the *a priori* information. $H(x)$ is the forward model, which maps the background information to a "background measurement", in our case, the refractivity, N.

$$y = H(x) \tag{4.2}$$

E and F are the expected covariances of the measurements and forward modeling, respectively. By solving (minimizing) this penalty equation using the least squares principle, profiles of pressure, temperature and water vapor can be obtained simultaneously. The gradient of the penalty function J is:

$$\nabla J = B^{-1}(x - x_b) - {H'}^T (E + F)^{-1}(y_0 - H(x)) \tag{4.3}$$

The penalty function (4.1) can be minimized iteratively using the steepest descent method, conjugate gradient method or quasi-Newton method. The covariance matrix Q of the final solution is given by [*Bouttier and Courtier*, 1999]:

$$Q^{-1} = J'' = \nabla\nabla J = B^{-1} + {H'}^T (E + F)^{-1} H' \tag{4.4}$$

In our simulation, the state vector consists of 33 elements:

$$x = [T(P_{16}), \cdots, T(P_1), \ln(Q(P_{16})), \cdots, \ln(Q(P_1)), P_s]^T \tag{4.5}$$

where T and Q are the temperature and the specific humidity on 16 pressure levels from 1000 hPa to 10 hPa. P_s is the surface pressure. The measurement from GPS occultation y_0 is a one-dimensional vertical profile of refractivity as a function of impact factor at fixed tangent points:

$$y_0 = [N(a_1), N(a_2), \cdots, N(a_{nobs})]^T \tag{4.6}$$

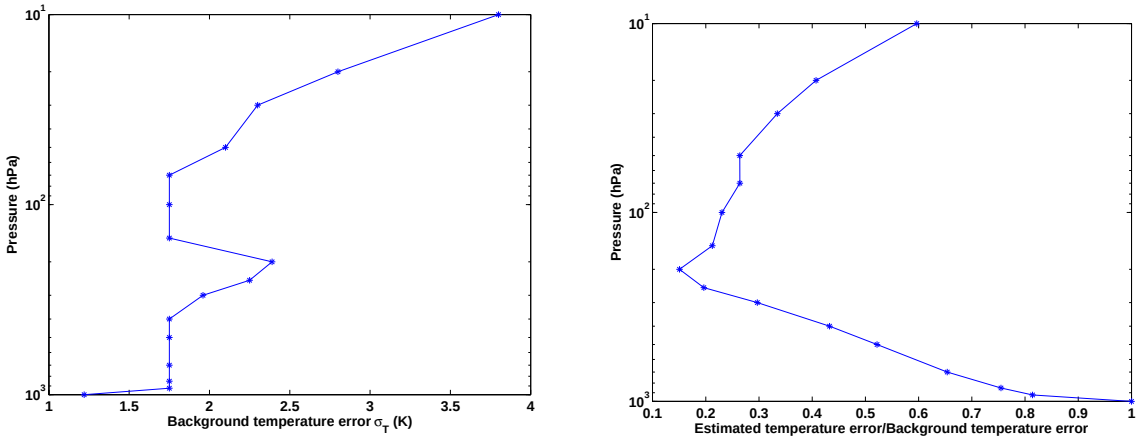
The background information is taken from ECMWF analysis. Its covariance matrix is constructed as diagonal for this experiment, since there is no full matrix available to use. The observation error matrix E is formed empirically as follow:

- The percentage error in refractivity is 1% at the surface and decreases linearly to 0.2% at the height of 10 km. Above 10 km, the error remains 0.2%.
- The correlations are assumed to be an exponential decay with the separation in geopotential heights, given by:

$$E_{nm} = \sigma_n \sigma_m \exp(-l(z_n - z_m)) \qquad l = 3 \times 10^{-4}\ m^{-1} \tag{4.7}$$

  where σ_n and σ_m are the observation errors at the height z_n and z_m.

The forward model $H(x) \rightarrow y$ converts the state vector x at model levels to the refactivity N at observation levels. Usually, the observation levels have much higher resolution than the model levels. First, the virtual temperatures and geopotential heights at model levels are calculated, which enables the calculation of the gradient of ln(specific humidity), temperature and virtual temperature with respect to the geopotential height between any two adjacent model levels. Then the specific humidity , temperature and pressure at any observation height between two adjacent model levels can be calculated using linear interpolations and gradient values. Finally, the refractivities are obtained at the observation levels. The forward modeling error F is ignored in our simulation by assuming that the interpolation errors from the model levels to the observation levels are minor. The gradient of the forward model H' is required by the gradient of the penalty function J in (4.3) and the covariance matrix of the final solution Q in (4.4). The procedure of calculating H' is simple but very tedious, one needs to apply the chain rule to the formulas in each step of the forward model calculation.



(a) (b)

Figure 4.13: The 1DVar simulation results.

Figure 4.13 (a) shows the simulated background errors for the temperature. After assimilating the GPS occultation measurements, the ratio between the estimated error and the background error of temperature shows great improvement on the 200-300 hPa region (Figure 4.13b). The errors are reduced up to 80%-90%. The error for the surface pressure obtained from the covariance matrix of the solution is reduced to 70% of the original background value (2.5 hPa). Since the surface pressure is related to the variables on each layer through the hydrostatic equation, any change of the observations on any layer could affect the surface pressure value. Consequently, the highly accurate GPS occultation observations at 5-30 km can bring positive impact on reducing the uncertainty of the surface pressure. The improvement of other state variables (e.g., temperature) can also indirectly improve the accuracy of the surface pressure. Although our simulation is only based on limited test, it does show that GPS occultation measurements have potential to improve the pressure modeling.

CHAPTER 5

THE EFFECTS OF ATMOSPHERIC PRESSURE ON TIME VARIABLE GRAVITY SOLUTIONS

5.1 The atmospheric pressure

5.1.1 The mean pressure field

Mass variations of the atmosphere contribute to the total mass change of the Earth system. The total pressure measurement (total surface pressure = dry air surface pressure + water vapor surface pressure) on the Earth surface (or integrated column of air) reflects the mass change in the atmosphere. The total mass of the atmosphere has been studied and constantly updated as the most recent model analyses become available [*Trenberth*, 1981; *Trenberth and Guillemot*, 1994; *Trenberth and Smith*, 2005]. The total mass of the atmosphere from a recent analysis *Trenberth and Smith* [2005] based on the 40-year ECMWF re-analysis (ERA 40) is 5.1480×10^{18} kg, and the water vapor changes annually in a range of 1.2 or 1.5×10^{15} kg depending on whether surface pressure or water vapor data is used. In his previous research [*Trenberth and Guillemot*, 1994], the total mass from the operational analysis is 5.1441×10^{18} kg with a change of 1.93×10^{15} kg throughout the year associated with the changes of water vapor in the atmosphere. According to his paper, the increase of the total mass is mainly due to the continuous improvement of the topography model.

It is usually assumed that the total mass of the Earth system is conserved. We do not solve the total mean mass (degree zero term in spherical harmonics) in gravity solutions. The approximate mass of the Earth is around 5.97×10^{24} kg and the mass of the atmosphere is only 10^{-6} of the total mass of the solid earth. The global mean surface pressure given by *Trenberth and Smith* [2005] is 985.5 hPa, and the water cycle has a maximum fluctuation in July of 2.62 hPa and a minimum in December of 2.33 hPa. Figure 5.1 shows a time series of the global mean surface pressure from the NCEP re-analysis.

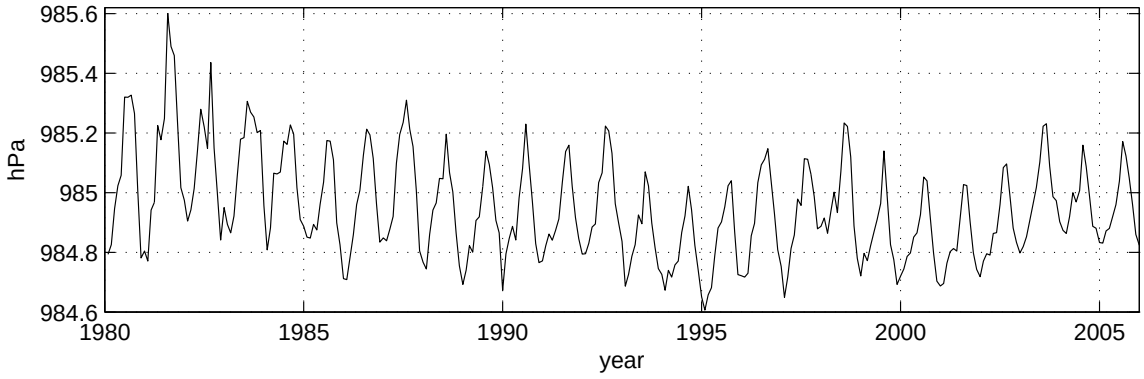


Figure 5.1: The global mean surface pressure from the NCEP re-analysis.

5.1.2 The variation of atmospheric pressure

Although the global mean atmospheric mass change is in a small range, the redistribution of the atmospheric mass is continuous. As one of the biggest signals in the time variable gravity field, atmospheric phenomena can be classified into two categories based on their time variability. The first category represents long-term atmospheric variations, including seasonal, annual, and inter-annual variations. Because of their long term characteristics and large magnitudes, these variations definitely can be sensed by satellite-to-satellite missions like GRACE. The features include the global atmospheric pressure annual cycle, which is high in July and low in December (see Figure 5.1), the seasonal variation between the northern and southern hemispheres, the formation of regular high-low pressure systems, and the mass movement between land and ocean, such as the Indian monsoon. As for the inter-annual variability, the El Niño-Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO) are the dominant features for the atmosphere. Most of these features can be resolved or represented by numerical models. The validation of such models requires multiple years of gravity observations. The other category reflects the short-term atmospheric variations. The most popular ones include the mid-latitude cyclones and anticyclones which can last for less than 10 days and cause 10-50 hPa surface pressure variation. The other short-term variations are the diurnal and semi-diurnal solar tides caused by heating and solar radiation [*Chapman and Lindzen*, 1970]. The atmospheric tides are more prominent in the tropics. The tides can be different by up to 2.0 hPa with respect to the mean daily pressure [*van den Dool et al.*, 1997].

The variation of the surface pressure is typically in the range of 10-20 hPa. The pressure difference or the pressure gradient force induces air flows or winds. In tropical

regions, winds are weak and the surface pressure anomaly is expected to be small. In the mid-latitude and polar regions, the surface pressure anomaly is larger. Figure 5.2 shows the RMS pressure variability using the NCEP re-analysis data in 2003. The mid-latitude region has the largest pressure variability. Some extreme conditions may also happen to introduce a large surface pressure anomaly. A hurricane is one of such phenomena that originates from the oceans. In the U.S., most of the hurricanes happen around the Gulf of Mexico region. A hurricane is an extremely low pressure system. The surface pressure at the center of the hurricane could drop 100 hPa in a very short period. The phenomena mentioned above may change rapidly, and may not be totally sensed by the gravity missions, but their resonance could affect certain orbit configurations and introduce observable secular or other long term perturbations. These high frequency signals also interfere with the gravity measurements, thereby they severely corrupt the lower frequency signals.

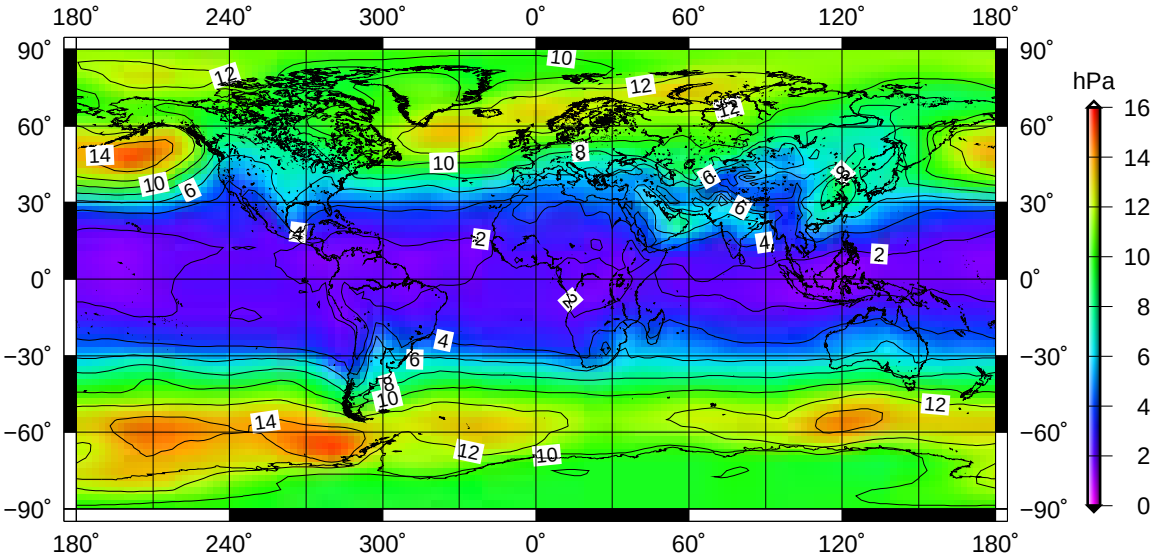


Figure 5.2: The RMS of global surface pressure for 2003 from NCEP re-analysis.

5.2 The pressure measurements and modeling

Pressure is the least variable property of the atmosphere compared to other properties. For temperature or precipitation, a small shift of the location could introduce a substantial error. Usually, there are two sources to obtain the pressure measurement, one is from *in-situ* measuring devices, such as barometers, anemometers, radiosondes

etc. Another way is to obtain data from numerical model analysis results. Real meteorological observations are not evenly distributed to cover the globe, but results from Numerical Weather Prediction (NWP) models provide spatially uniform coverage and internally consistent datasets. The most commonly used NWP models are those from ECMWF and NCEP. These models are used to predict the future state of the atmosphere by modeling the dynamic and physical processes of the atmosphere. Modern numerical forecast models are based on a formulation of the dynamic equations, which are referred to as primitive equations. Variables in the models are represented by either grid or spectral forms. Currently, most of the global operational models utilize spectral models for better computational efficiency. Grid models are generally adopted in high resolution, regional or limited-area models (e.g. MM5). To integrate the dynamic equations, accurate initial conditions are required to be consistent with the model initialization. This dataset is derived from various observations, including ground weather stations, upper air stations (e.g., radiosonde), ship, aircraft and polar and geostationary satellites. The process of establishing the grids for the initial condition from all the observations in a limited time span is called data assimilation. The processes include objective analysis, which is the error checking and the interpolation of the observation to the grids, and data initialization, which removes the spurious wave noise to prevent them from exaggeration. The resulted analyzed grid represents a best approximation (estimation) of the true state of the atmosphere at a given time, which is a linear combination of the observations and a first guess from previous forecasts. We mainly use the analyzed (not forecasting) fields in this study.

Two kinds of analysis products are usually available: operational analyses and reanalyses. The assimilation method of operational analyses is constantly evolving as time goes on. Variations of the initialization methods, physics, grid resolution, data availability, and handling techniques will disrupt the consistency and the continuity of the analyses. To provide a consistent dataset, many organizations established their own project to produce the re-analyzed fields for various time spans such as the ECMWF 40-year re-analysis (ERA40) and NCEP reanalysis. In these products, the assimilation schemes do not change over time, but the observations will change due to the data availability over time (e.g., no satellite data in the past). Unfortunately, the ECMWF 40 year re-analysis does not cover the time span for GRACE. Instead, ECMWF operational analyses will be used.

5.3 Datasets from analyses models

The analysis data we use are from ECMWF WCRP/TOGA data sets (TOGA stands for Tropical Oceans Global Atmosphere, a program under the World Climate Research Program (WCRP)) and the NCEP operational analyses. The ECMWF archive is directly interpolated from the ECMWF operational, full resolution, surface, and model level data. The dataset is represented in a horizontal resolution of T106

(truncated at wavenumber 106) in spherical harmonics, and vertical resolution of 21 standard pressure levels from 1000 hPa to 1 hPa, although the original models run at a much higher resolution (T799 in horizontal and L91 in vertical as of 2006). Geopotential height, temperature and relative humidity are available 4 time daily at 00Z, 06Z, 12Z, 18Z (Z stands for Greenwich time).

The NCEP operational dataset we use are those from the Final (FNL) Global Data Assimilation System (GDAS) on a $1^\circ \times 1^\circ$ grid covering the entire Earth every 6 hours. It is one of the models from the NCEP's global parallel system. The other alternative is from the Global Forecast System (GFS). The FNL run provides the best analysis results since it waits to collect data for at least 6 hours past the synoptic time, while GFS only incorporates data for 2 hours 45 minutes past the synoptic time. The FNL runs at the highest resolution T254 and L64, which corresponds to 55 km horizontal grid. GFS runs at the same resolution but with less data included. We choose FNL over GFS in the study and it is on a regular latitude-longitude grid. In the datasets, geopotential height, temperature, and relative humidity are available on surface and 26 mandatory pressure levels from 1000 hPa to 10 hPa. Both datasets were obtained from the National Center for Atmospheric Research (NCAR).

Both datasets are stored in GRIdded Binary (GRIB) format, which is a widely used bit-oriented format proposed by the World Meteorology Organization (WMO) to exchange large volume of data among the meteorological communities. Analyses centers do not strictly follow the standard, and format variations exist for different products. The description for the NCEP dataset can be found from the NCEP GRIB document ON388 [*Gordon*, 2002]. The local GRIB extensions for ECMWF datasets are listed at the official ECMWF website.

5.4 The effect of topography on surface pressure

For *in-situ* barometric measurements, the use of the data is straightforward, regardless of the instrument measurement error. Some NWP analysis products directly include the surface pressure. However, most of the time, these values are not suitable to be used directly. The terrain in numerical weather models (or model orography) are converted from Digital Elevation Models (DEM) using different methods, which include (1) mean orography (average of the terrain data inside a model grid box); (2) envelope orography (cover all but the very sharpest peaks); (3) Silhouette orography (averages only the tallest features in each grid box). Comparison of the surface pressure directly from model output may cause some problematic effects. Since the topography alters the flow of the atmosphere, the real topography is adjusted or smoothed to best fit the physical model and its resolution. Sometimes, biases exist between the real and the artificial model surface pressure, i.e., the surface pressure from the NWP model is different from the surface pressure on the real physical Earth. The accuracy of the conversion from the 3D global analysis data to surface pressure

is greatly affected by the complexity of the land topography. The surface topography represented in the models also differs by its resolution depending on the scale of the model. In the region with the steepest elevation change, the difference and variability are more obvious. In addition, many NWP models use surface spherical functions to represent the surface topography, which may lead to spurious ripples in the area of steep elevation change. This is the Gibbs phenomenon caused by the truncation of the spherical harmonics. The topographical discrepancy in coastal areas is also largely due to the adoption of different land sea masks in the models. The procedures for the computation of the surface pressure is briefly introduced in appendix B.

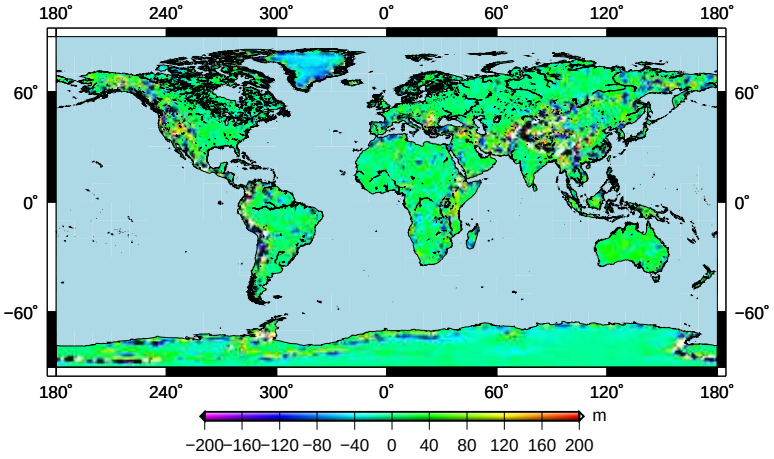
Figure 5.3 shows the differences among ECMWF, NCEP, and ETOPO2 topography. ETOPO2 is the 2-minute Gridded Global Relief data from the NOAA National Geophysical Data Center. This topographic mosaic is converted to a grid resolution of $1^\circ \times 1^\circ$. In the global and Antarctic maps, significant differences are found in the high elevation areas and along the coastal lines. ECMWF and NCEP also use different elevation data in their models. Therefore, we cannot trust the surface pressure data in these areas.

In the GRACE de-aliasing model (see chapter 7), a mean field covering at least one year is removed from the 6-hourly model output. Most of the effects of the topography in the residual atmosphere are canceled. The remaining part is not on the same order of magnitude.

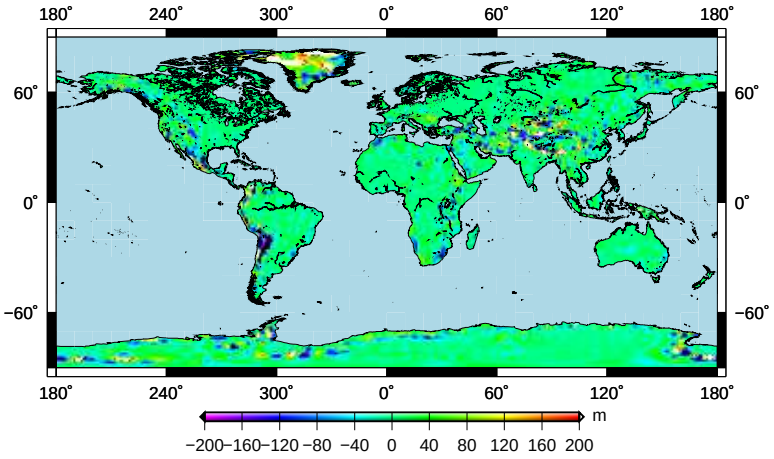
5.5 The uncertainties of the NWP surface pressure

In the GRACE processing procedures, atmospheric effects need to be first removed before solving for the time variable gravity signal. In order to evaluate the NWP model performance, *Velicogna et al.* [2001] compared the surface pressure from ECMWF and NCEP with surface barometric measurements in the United States and on the Arabic peninsula. In their analysis, the point accuracy of the global model is around 1.0-1.5 hPa. The accuracy could be improved to 0.5 hPa by taking the spatial and temporal average over a period of one month. The authors concluded that the analyzed pressure fields are adequate to remove the atmospheric contribution from GRACE hydrological estimates to sub-centimeter levels.

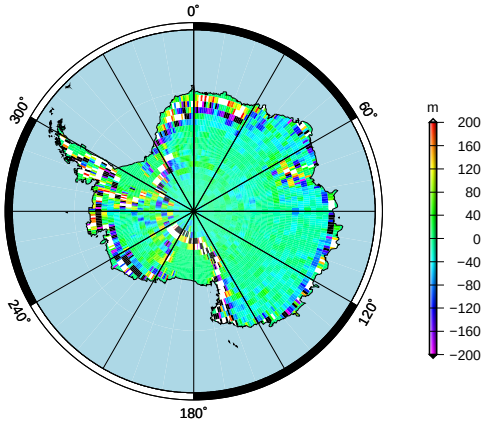
However, this analysis does not apply to the Southern Ocean and Antarctica, where models have a bad performance because of lack of *in-situ* measurements. Models are also not consistent with each other. To compare the surface pressure, we interpolate the ECMWF and NCEP pressure fields to the same topographical surface. Figure 5.4 (a–e) shows a 4 times daily example of the pressure difference between NECP and ECMWF based on the ECMWF topography. Only the 00Z difference is plotted as a global view in (a). Significant difference (8 hPa) can be found over some regions in the Southern Ocean and the Antarctic continent. The pattern alters with time. This large disagreement between the two models shows the uncertainty and



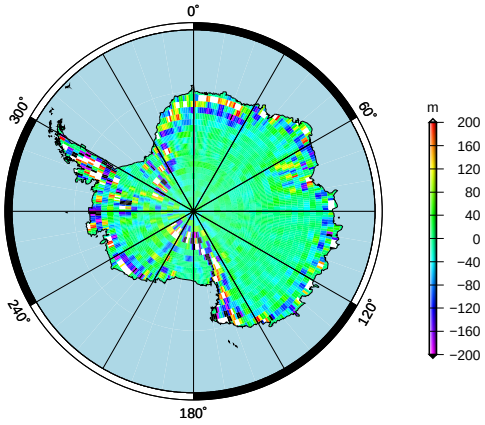
(a) ECMWF-ETOPO2



(b) ECMWF-NCEP



(c) ECMWF-ETOPO2



(d) ECMWF-NCEP

Figure 5.3: The differences of ECMWF, NCEP and ETOPO2 topography.

inconsistency of these two data products over the Southern Ocean and Antarctica. We also examine the difference in the monthly mean field in Figure 5.5, where the ECMWF and NCEP agrees reasonably well except for the southern high latitude region. A 6 hPa discrepancy is found in the east Antarctic region.

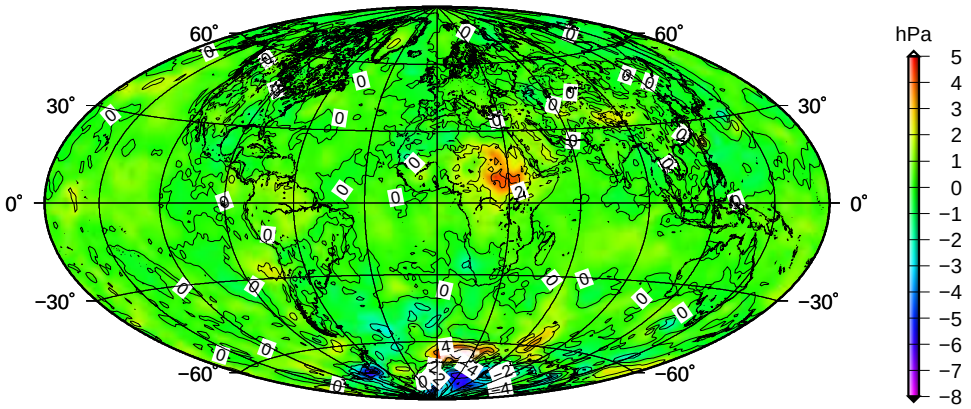
Compared to other regions of the globe, the Southern Ocean and Antarctica are data-sparse regions. The lack of observation networks and traditional measurements greatly affects the model performance. We can not tell which model performs better by simply comparing them. *In-situ* observations are needed to validate their respective performances. However, a large part of these observations (e.g., surface pressure) have already been assimilated into these models, which means that the model result is a weighted sum of the real observation and a background model, and the observations are not fully independent. Meanwhile, ECMWF and NCEP also have errors in common. If we use such datasets, the comparison results would be a little optimistic. As discussed earlier, the analyzed fields do not necessarily represent the true state of the atmosphere, they are mainly used as an initial condition for forecasting purposes. The initial guess field (background field) may have a weight as much as the real observations. In Antarctica, the *in-situ* observations are very limited, we have to neglect this effect in the comparison, but we should carefully interpret these results.

5.6 Effects of atmospheric pressure on the GRACE gravity solution

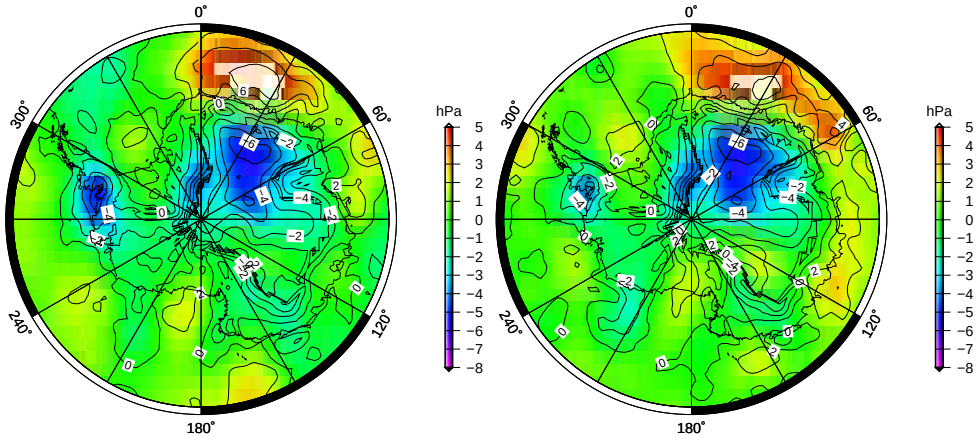
By measuring with space geodetic techniques, signals are contaminated when passing through the Earth's atmospheric envelope. The atmosphere can delay the distance measurement, cause deformation of the surface of the solid Earth, and redistribute mass as air flows.

The main objective of the GRACE mission is to recover the time variable gravity field signals with a resolution of several hundred kilometers or larger, and monthly time resolution accuracy around 2 mm of water thickness over land and 0.1 hPa or better for ocean bottom pressure. Such a signal can constrain the processes of mass redistribution and facilitate the research related to the geophysical evolution and circulations on the Earth.

The atmosphere error also affects the determination of the non-steric components (caused by the change of mass) in the sea surface height using GRACE and altimeter measurements. The mass variation causes changes in both GRACE ocean bottom pressure measurements and altimeter measurements, while the steric components (change of density) are only reflected at the altimeter sea surface height measurements. The combination of the two can determine the non-steric part of the sea surface height change. There are two approaches to remove the mass redistribution from the atmosphere above the ocean as suggested by *Wahr et al.* [1998]. One is to remove the mass of atmospheric redistribution without removing the ocean response

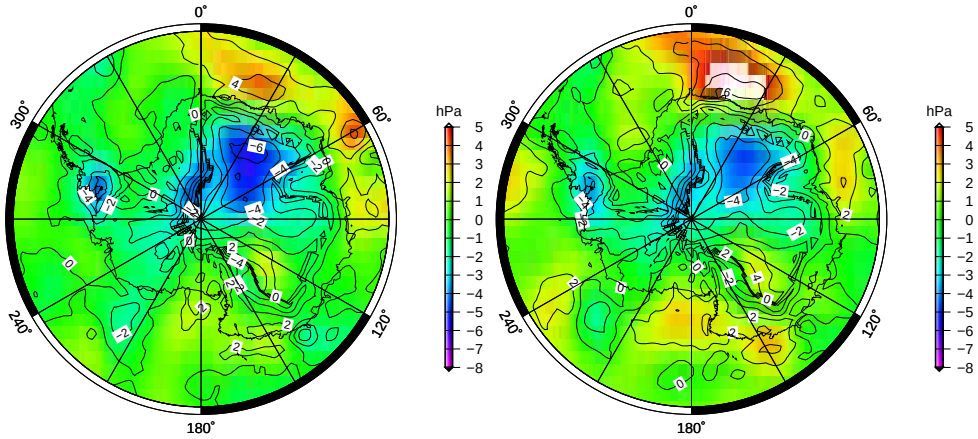


(a) July 01, 2004,00Z, ECMWF-NCEP



(b) July 01, 2004, 00Z

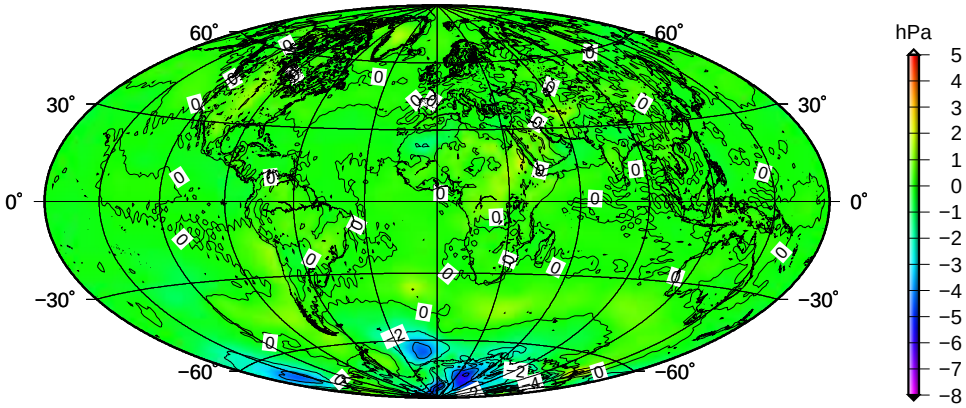
(c) July 01, 2004, 06Z



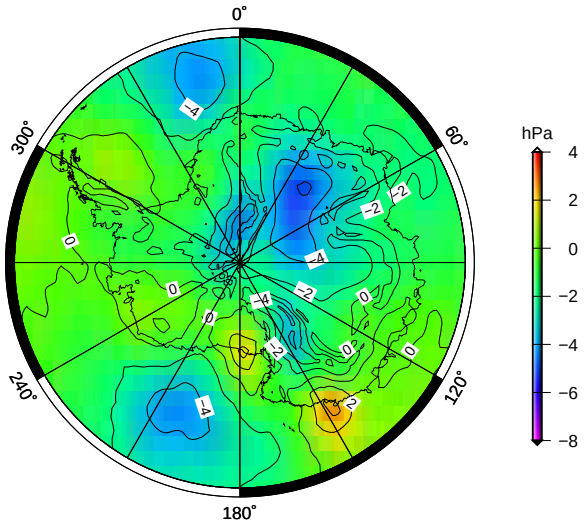
(d) July 01, 2004, 12Z

(e) July 01, 2004, 18Z

Figure 5.4: The differences between the ECMWF and NCEP surface pressure on the ECMWF topography.



(a) July 2001, Monthly difference: ECMWF-NCEP



(b) July 2001, Monthly difference: ECMWF-NCEP

Figure 5.5: Monthly surface pressure difference between ECMWF and NCEP on the ECMWF topography.

to such redistribution. The error in the atmospheric product contaminates the gravity measurements (ocean bottom pressure). The other is to remove both the atmospheric mass redistribution and oceanic response from GRACE data, as well as the ocean's equilibrium response to the atmosphere from the altimeter measurements. In this case, atmospheric errors enter into the altimeter measurements.

Hydrological information is also greatly affected by the accuracy of the surface pressure measurement. The uncertainty of surface pressure errors over the land area varies vastly depending on the region of interest. The Southern Ocean and Antarctica have the highest uncertainty concerning the meteorological measurements.

The uncertainty of mean surface pressure, particularly at Antarctica, is a big problem for retrieving the post glacial rebound signal. This has implications as to the deterioration of the ice-sheet and mass balance study. The sparsely located stations have a great impact on the performance of the numerical models over such an area. The inter-comparison of two popular meteorological models (ECMWF and NCEP) shows some inconsistency, the difference between them usually exceeds 6 hPa.

5.6.1 Atmospheric errors in GRACE solutions

The atmospheric mass signal is relative large. Figure 5.6 shows that signals from ECMWF surface pressures are bigger than the GRACE GGM02S [*Tapley et al.*, 2005] estimated (approximately calibrated) error at around degree 30. The difference between ECMWF and NCEP is below the current sensitivity, but is larger than the pre-launch sensitivity (40 times better than the current sensitivity) at around degree 30. The error curve for GRACE includes the contributions from the un-modeled atmospheric signals and measurement errors. Those errors degrade the accuracy of the recovered gravity signals.

One of the GRACE science products is the monthly gravity solution, which is a set of spherical harmonics computed from one month of GRACE observations. From these piecewise monthly solutions, one can infer the mass variation. The residual pressure variation is also included in this solution. Since we are only interested in the variational signals, the total pressure is separated into two parts:

1. A long-term mean surface pressure

2. Short period pressure variations

The long term mean surface pressure is included in the static gravity field, while the remaining short period variation is reflected in the monthly solution. There might be errors in this mean surface pressure, but it is easy to fix once a better mean field is available. Meanwhile, if we subtract one monthly solution from another, the common errors are canceled. Among different mass variations contained in the monthly solution, atmospheric variations are better known from numerical models than other signals (e.g., hydrology). To separate the total mass variation, atmospheric mass

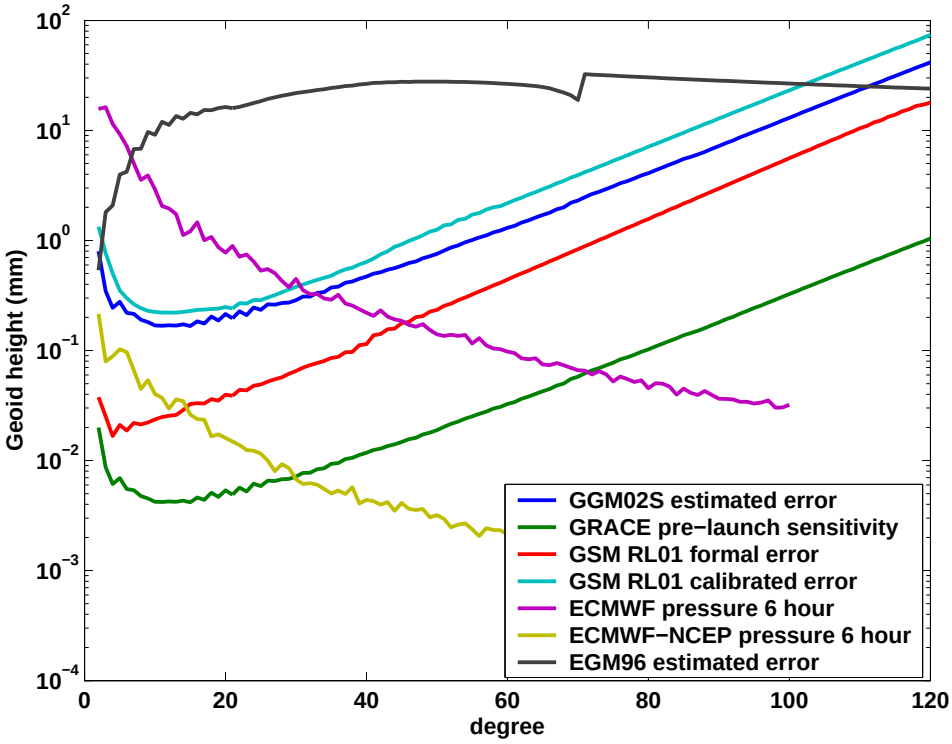


Figure 5.6: The sensitivity of GRACE to atmospheric surface pressure.

variations need to be accurately removed. Outputs from the numerical meteorological models contain errors. These errors enter into a monthly solution in two different ways:

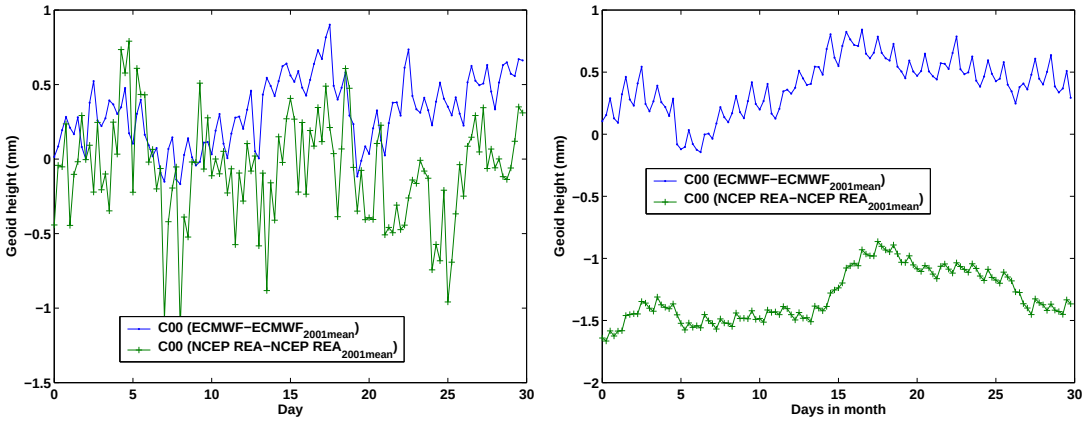
1. The mis-modeled and un-modeled pressure errors
2. The un-modeled pressure variability – aliasing error

The mis-modeled and un-modeled pressure errors are those signals that can be correctly sensed by GRACE, but are missing or mis-represented in weather models. The GRACE solution can capture most of the large spatial features represented by low degree spherical harmonics. However, because of the continuous motion of the large atmospheric mass during the estimation span (30 days), part of the signals (e.g., at the boundary) are interpreted as short wavelength features. As near polar-orbiting satellites, the twin GRACE satellites can achieve a denser spatial sampling while there is no improvement for the temporal sampling. The high-degree features which mostly consist of the short wavelength signals can not be fully captured by GRACE. The aliasing error mainly comes from the orbital characteristics of the satellites and

the mis-representations of the variability by the NWP models. If de-aliasing (nominal) models can correctly capture the true variability of the atmosphere, the monthly gravity estimations can produce the mean difference between the true atmosphere variation and the nominal model. Otherwise, solutions are corrupted by aliasing, which will be discussed in detail in later sections.

5.6.2 The degree 0 and 1 terms

The pressure field is usually transformed to spherical harmonics before it is used as a GRACE background model. The degree 0 and 1 terms are excluded from the later computations. These terms can affect some of our later analyses. We should mention these factors here. The degree 0 term reflects (is proportional to) the total mass of the atmosphere. For the whole Earth system, this term represents the total mass of the Earth, which is often assumed to be a constant. While for each subsystem, the total mass is not necessarily zero, the net mass variation of the whole system is zero. Therefore, the $\Delta C_{0,0}$ term can be neglected from the GRACE solutions. Figure 5.7 shows the change of this term in two one-month periods. They do not agree very well, which implies that the NWP models generate different mass fluxes to other subsystems. The bias in the 2003 NCEP operational analysis may indicate some model change with respect to the pressure field in 2001.



(a) (b)

Figure 5.7: The $C_{0,0}$ term calculated from surface pressure models, a 2001 mean field is removed from each time series. (a) ECMWF and NCEP reanalysis in April, 2001; (b) ECMWF and NCEP operational analysis in April 2003.

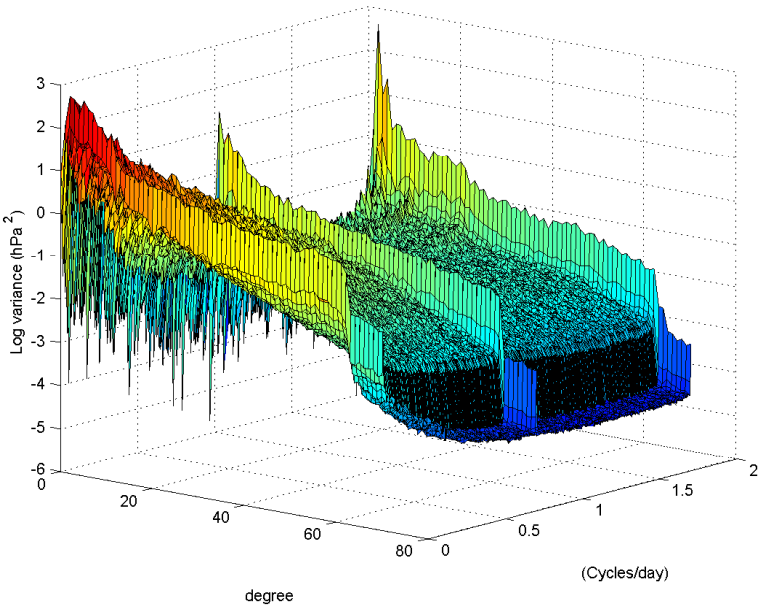
The degree 1 terms represent the relative position between the Earth's center of mass and the center of the coordinate system. For the mass change in the atmosphere, these terms are also non-zero. In the GRACE data processing, the coordinate center is assumed to coincide with the center of mass. Thus, the degree 1 terms also vanish.

The RMS value is often used to indicate the variation of the sampled data. When computing the RMS value of the surface pressure, the direct use of the gridded values is inappropriate. Degree 0 and 1 terms must be removed first. The statistics with and without the degree 0 and 1 terms are different, which might affect the interpretation of the variability of the surface pressure change.

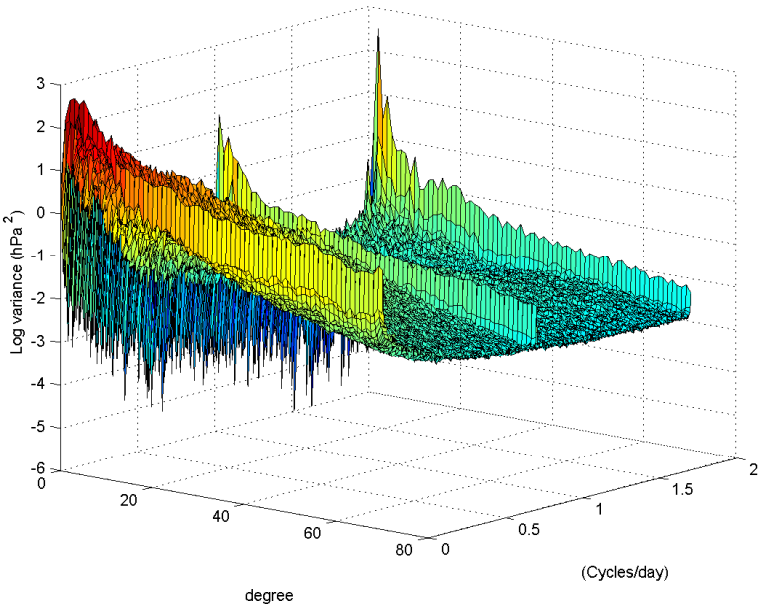
5.6.3 The degree 2 terms–barometric tides

As mentioned earlier, the amplitude of the barometric tides caused by the thermal forcing can be as large as 2 hPa. It is relatively small compared to other variations, but the signal is still sensible to GRACE. The most dominant barometric tides are diurnal (S_1) and semi-diurnal (S_2) tides, with a larger amplitude for semi-diurnal tide. Analysis products from meteorological centers are usually given every 6 hours, which just coincides with the Nyquist frequency of the S_2 tide. Consequently, the S_2 tide waves resolved from these products are stationary rather than moving westward as predicted by theory. The sampling rate can also result in reduced tide amplitude if the tide does not happen right at the sampling time. To recover the traveling S_2 tide, *van den Dool et al.* [1997] proposed a temporal interpolation method to resolve the temporal change of the tide phase.

Figure 5.8 shows a frequency-degree variance spectrum for the surface pressure fluctuation in the NCEP and ECMWF analysis fields using one year of 4 times daily data from each dataset. The method is identical with that described by *Wunsch and Stammer* [1995] for altimetric datasets. In the figure, there are two sharp spikes around 1 cycle/day (cpd) and 2 cpd (see Figure 5.9), which indicate that the energy for S_1 and S_2 are dominant. Following the interpolation method by *van den Dool et al.* [1997], some researchers examined the S_2 tide derived from gridded data using NCEP re-analysis and ECMWF operational analysis against the meteorological station data [*Ray*, 2001; *Ray and Ponte*, 2003]. They found that the ECMWF derived tide is more realistic than the tide derived from other commonly available surface pressure datasets after properly correcting the systematic phase error. Most of the ocean models account for the solar radiation tide, and the air tides are usually treated as part of the ocean tide. Therefore, before forcing the ocean model, the S_2 atmospheric tide needs to be filtered out from the surface pressure data. A procedure was proposed by *Ponte and Ray* [2002] based on the algorithm of *van den Dool et al.* [1997]. In the GRACE data processing, the S_2 tide needs to be first removed from the $\bar{C}_{22}$ and $\bar{S}_{22}$ of the Atmosphere and Ocean De-aliasing 1B (AOD1B) product. The value of S_2 is obtained from the difference between the TEG4 estimation of this tidal harmonic



(a) NCEP re-analysis from year 2003



(b) ECMWF operational analysis from year 2001

Figure 5.8: Frequency-degree variance spectrum for surface pressures from NCEP re-analysis and ECMWF operational analysis.

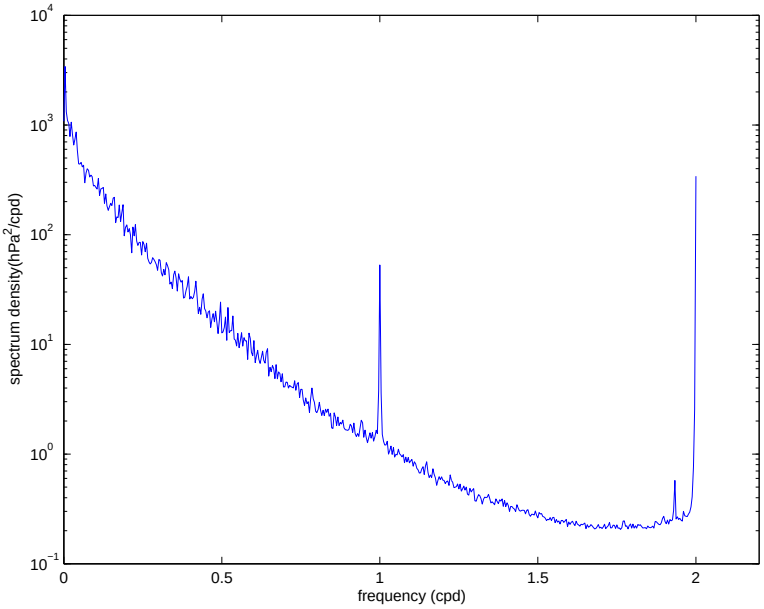


Figure 5.9: Globally integrated spectrum of ECMWF surface pressure for year 2001.

and altimetric determination of the tidal harmonic from the CSR 4.0 tide model [*Bettadpur*, 2004]. In this way, the ocean tide model will include the S_2 atmospheric tide.

5.6.4 Inverted barometer effect

Under the Inverted Barometer (IB) assumption, the atmospheric error over the ocean is ignored. This is because that we assume an equilibrium response of the ocean to the pressure [*Wunsch and Stammer*, 1997]. An equilibrium response is very close to the IB response, i.e, 1 hPa increase in atmospheric pressure depresses the ocean surface by 1 cm. There is no net change of ocean bottom pressure. In other words, there is no mass change when integrating the vertically overlaying atmosphere and the ocean column. Usually, ocean does not respond like IB perfectly, even with a longer average time for GRACE. Previous study by *Ponte and Gaspar* [1999] found significant deviations from IB in the tropics and some regions of the Southern Ocean. Wind effects can explain most of the deviations in the middle and high latitudes, and true non-IB effects are responsible for most of the decreases in the tropics. Therefore, both IB and non-IB are incorrect, the real response may reside somewhere in between. Under the IB assumption, atmospheric variations are required to be removed only over land. The error source includes the atmosphere error over the land area and

the leakage of the atmospheric signal along the coasts through an averaging function. In the following, we demonstrates the aliasing error under both the IB and non-IB assumptions.

5.6.5 The atmospheric aliasing

Aliasing is caused by the fact that the sampling frequency is lower than the highest frequency of the signal. Polar orbiting satellites can nearly cover the entire globe within a certain period of time. However, the data measured in this way are both time and position dependent. Depending on the orbital design, some areas of the earth are densely covered, while other areas are under-sampled such as the gap between two satellite tracks. Temporally, the observations covering the globe are not synoptic (observed at the same time) and the time interval for subsequent observations surrounding the same location is usually not a constant.

The time-variable product from GRACE is usually the monthly solution. Short period variations of the atmosphere, ocean, and other non-tidal signals are merged into the monthly solution due to the space and time aliasing. Among those signals, the atmosphere has the biggest input. Therefore, de-aliasing models are critically needed. Unlike the tidal signal, the spectrum of the atmosphere signal is very broad, which makes the effect of aliasing on the gravity solution very complicated and hard to remove.

The best way to reduce aliasing is to use state-of-the-art weather products. *Han et al.* [2004] simulated the effect of the model error on the monthly GRACE solution using the energy conservation method [*Jekeli*, 1999]. It is shown that the unaccounted atmosphere corrupts the whole spectrum of the gravity solution and errors beyond degree 30 have larger magnitudes than the observation noise. *Thompson et al.* [2004] examined the impact of the short-period, temporal mass variation on the GRACE monthly estimate. It is found that the degree error relative to the measurement error increases by a factor of 20, due to the atmosphere aliasing; and the de-aliasing gives the greatest reduction in aliasing errors for the coefficients in mid-degrees and higher.

Since there is no easy way to quantify the real aliasing error in the GRACE solution, we exploit a similar simulation strategy as described in *Han et al.* [2004] and *Thompson et al.* [2004], but with alternative analysis models. Our simulation shows similar results as in previous studies, but there are also noticeable differences. The correlation between the RMS of pressure fields and the aliasing are examined under both IB and non-IB assumptions. In addition, we further demonstrate that the aliasing errors for the middle to high degree coefficients are mainly caused by the variability of the long wavelength terms. The effect of the interpolation of the 6-hour coefficients to the aliasing error is investigated as well. In the simulation, the aliasing error in Antarctica is also studies. To mimic the GRACE de-aliasing process, a mean field is pre-removed from each time series using its own mean in 2001. The differences

between the mean fields for different models are neglected. Chapter 7 describes the atmospheric de-aliasing model in detail.

5.6.5.1 Definition of the aliasing error

The aliasing error is defined by equation (5) in *Thompson et al.* [2004] and is given here for reference.

$$\epsilon = \delta\hat{G}(T_s) - (< \delta G_{true}(t) > - < \delta G_{nom}(t) >) \tag{5.1}$$

ϵ is the aliasing error. δG_{true} is the true perturbation potential, and $\delta G_{nominal}$ is the nominal perturbation potential. $<>$ denotes the time average. $\delta\hat{G}(T_s)$ is the updated gravity estimation during the time span T_s. In our simulation, ECMWF is used as the true model and two NCEP analyses are treated as the nominal models.

5.6.5.2 Simulation procedures

A 30-day orbit is generated by giving the GRACE orbital configuration as the initial condition. EGM 96 is used as the gravity model for the orbit. The inclination is 89°, and the separation of the two satellites is 2° at 450 km altitude. The integration step and the data output are both at 10 second interval. An error is applied to the observation to simulate the GRACE measurement error. The geopotential coefficients are inverted using a rigorous formula implemented in a parallel platform (32 node Linux cluster) [*Xie et al.*, 2004]. The coefficients are limited to degree 60 in both forward model and inversion process. To quantify the aliasing error in different scenarios, ECMWF TOGA is used as the true model both in April, 2001 and April, 2003. The nominal models are NCEP reanalysis for April, 2001 and NCEP operational analysis for April, 2003, respectively. The analysis models are only available every 6 hours. The simulation needs pressure values every 10 seconds. There are two methods to handle this problem. One is to treat the pressure as a constant within ±3 hours of the sample time, the other is to linearly interpolate the surface pressure within the 6 hour interval. Both cases are tested by examining their aliasing errors defined by the equation (5.1). It is found that the linear interpolation method has lower aliasing errors for low degree terms (Figure 5.10). As a result, we adopt the linear interpolation method in the simulation.

5.6.5.3 Results

Figure 5.11 shows the simulation results for 3 cases: IB and non-IB cases in April 2001 using NCEP reanalysis model as the de-aliasing model (top and middle), and non-IB case in April, 2003 using NCEP operational model as the nominal model (bottom). In each figure of the left panel, the quantities represented by the right side of the equation (5.1) are plotted by letting $\delta G_{nom}(t)$ equal to 0 or NCEP model

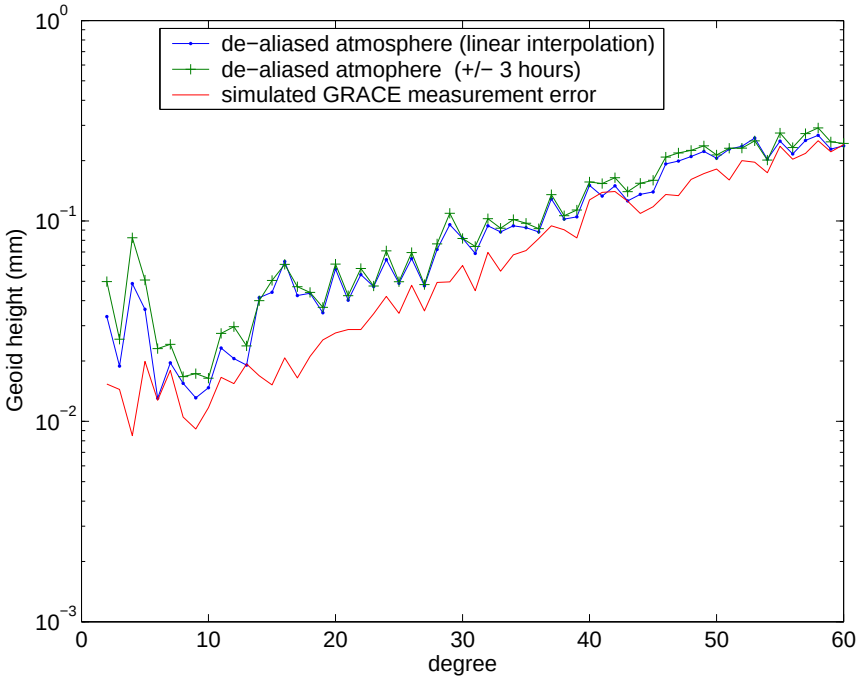


Figure 5.10: Comparison of two methods of using the surface pressure.

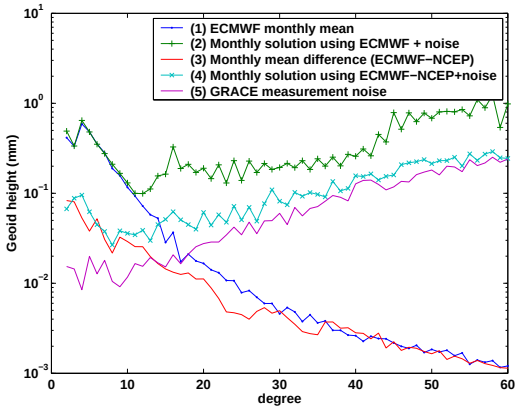
value to simulate the aliased and de-aliased cases. The simulated measurement error is plotted in all figures to indicate that it is an approachable, but unachievable ideal error. Curve (1) and (2) in (a), (c), and (e) of Figure 5.11 illustrate the monthly mean ECMWF field and the GRACE recovery result in the case of no de-aliasing model is present. The difference between curve (1) and (2) reflects the aliasing error without a de-aliasing model. In a similar way, the difference between curve (3) and (4) reflects the aliasing error with a nominal de-aliasing model. From the figures, we can see that the GRACE solution can capture most of the low degree ($< 10°$, long wavelength) features, though the agreement is not exact. The two curves start to deviate from each other from around degree 10, while a previous study found that the deviation occurred at around degree 30 [*Han et al.*, 2004]. (b), (d), and (f) in Figure 5.11 depict the aliasing error in each case. The most noticeable change after applying the de-aliasing model is the error reduction for the middle to high degree coefficients, which agrees with the previous study [*Thompson et al.*, 2004]. We also plot the calibrated error for the GRACE RL01 solutions, which can be treated as an upper bound for the GRACE error. The aliasing error is an optimistic estimate, since there are common errors in the ECMWF and NCEP fields. It is fair to guess that the real GRACE error resides between these two curves.

By comparing the curve (2) in figure (a) and (c), we can see that under the IB assumption, the mean field becomes much smaller. This is also confirmed by the geographic map (Figure 5.12). The aliasing error in the non-IB case is much smaller than in the IB case (Figure 5.11d). Some of the variability above the ocean is removed with the IB assumption. In some regions of the curve, the aliasing error approaches the measurement noise. For the simulation in April, 2003 using the NCEP operational analysis, it is surprising to see that the aliasing error becomes small even in the non-IB case (Figure 5.11f). It does not necessarily mean that the real aliasing error in the GRACE solution is small. It only indicates that the ECMWF and NCEP operational results are becoming more similar to each other, and NCEP can follow the variability of ECMWF very well. People commonly consider that ECMWF has a better performance. Thus our results implies that the NCEP operational model is continuously improving. Using variations of their difference to quantify the aliasing error becomes too optimistic when the two models become more similar.

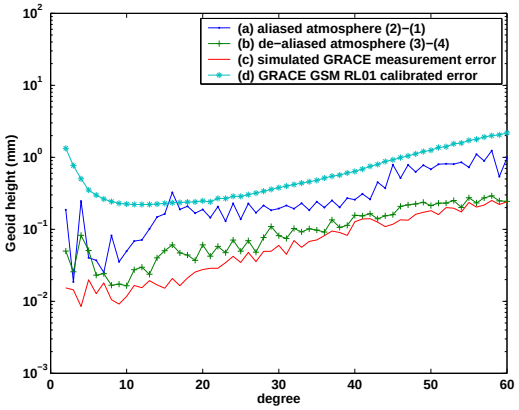
The monthly mean fields in Figure 5.12 still show big differences in the Antarctic continent. In the non-IB case (3 figures on the left), we can see large mass features around the Southern Ocean area. Such features are masked out in the IB case (3 figures on the right). The RMS value of the monthly pressure field can show us the variations of the pressure within one month. Large variations in the middle to high latitudes are shown on the RMS figure of the ECMWF field (Figure 5.13a). After the de-aliasing process, the RMS value becomes much (about 4 times) smaller (Figure 5.13c). The IB assumption omits the large variations in the Southern Ocean (Figure 5.13d). The NCEP operational analysis agrees with ECMWF very well, even in the non-IB situation, the RMS value is very small (Figure 5.13b). The small RMS value may indicate a possible model improvement.

The degree RMS does not give us a sense of the spatial distribution of the aliasing error. In Figure 5.14, the error in the recovered GRACE signal is compared with the truth error. Figure 5.14a shows the total errors embedded in the GRACE signal, including the aliasing error plus the mis-modeled or un-modeled error. After applying the Gaussian filter of 600 km radius (Figure 5.14b), the magnitude of the aliasing error is greatly reduced, the mis-modeled long wavelength error shows up. Figure 5.14c and Figure 5.14d illustrate the total aliasing error before and after applying the 600 km Gaussian filter. The biggest error is reduced by a factor of 8. Figure 5.14e is the true atmospheric error. Its difference with Figure 5.14b reveals the difference between the true error in the model and the total error in the GRACE solution. Obviously, the amplitude of some long wavelength errors is reduced by the smoothing. For example, the error over Antarctica in Figure 5.14b was reduced. One side effect of smoothing is that the amplitude of signals is also reduced.

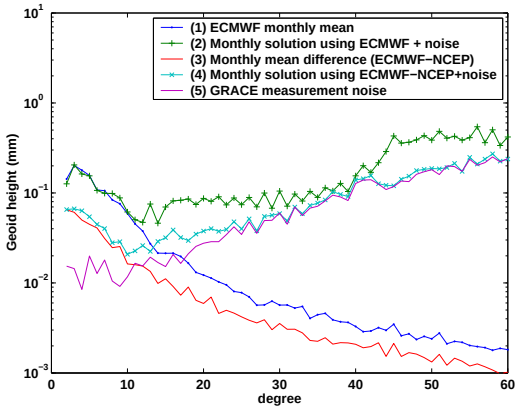
The aliasing effect is not a localized feature. We consider a situation that the pressure variations only exist around the Antarctic area (defined here as latitude $>$ 60°S). In all other areas, variations are set to zero. The simulation result (Figure



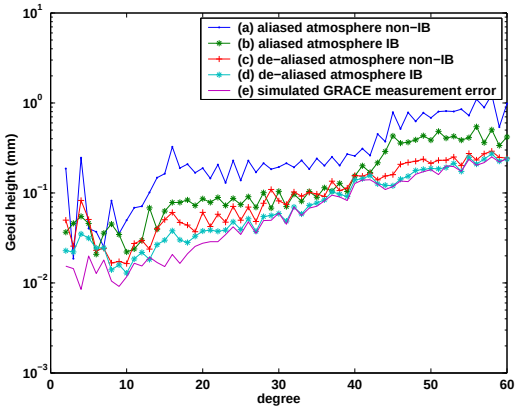
(a) Mean and recovered coefficients, non-IB, in April, 2001, NCEP reanalysis



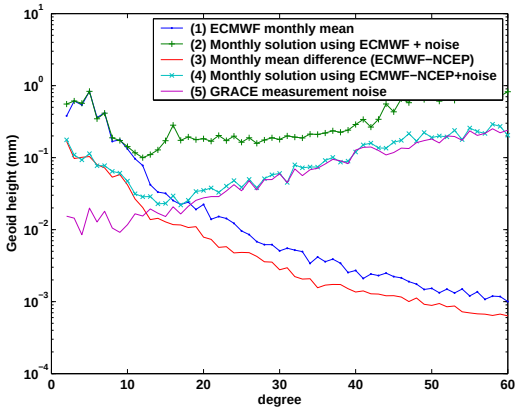
(b) The aliasing error, non-IB, in April, 2001, NCEP reanalysis



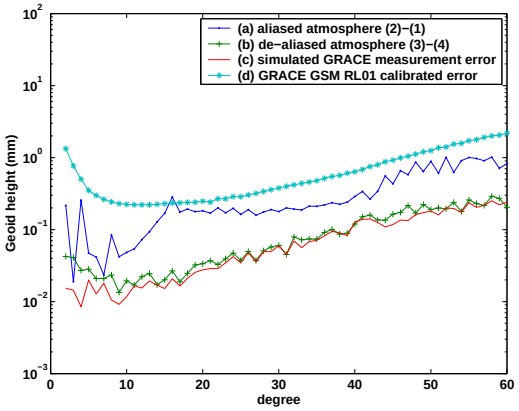
(c) Mean and recovered coefficients, IB, in April, 2001, NCEP reanalysis



(d) The aliasing error, IB, in April, 2001, NCEP reanalysis

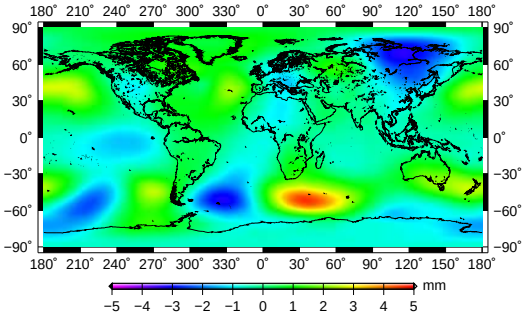


(e) Mean and recovered coefficients, non-IB, in April, 2003, NCEP operational analysis

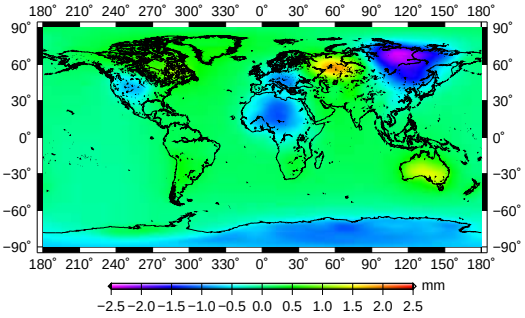


(f) The aliasing error, non-IB, in April, 2003, NCEP operational analysis

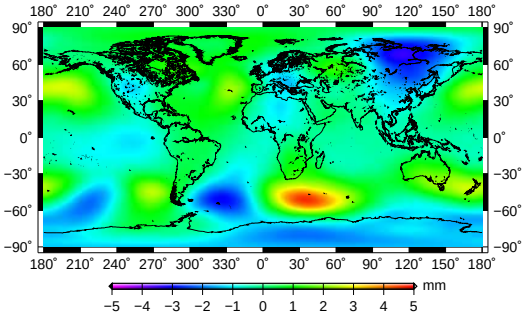
Figure 5.11: The degree RMS errors from atmospheric aliasing. Top: non-IB (April, 2001); Middle: IB (April, 2001); Bottom: non-IB (April 2003).



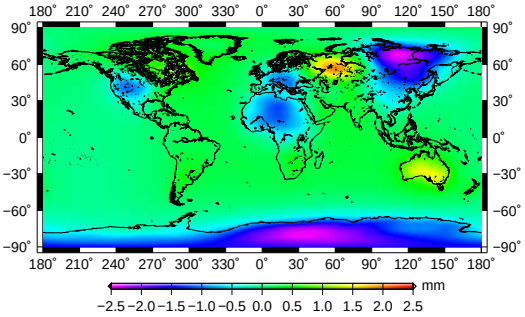
(a) Mean ECMWF surface pressure, non-IB, in April, 2001



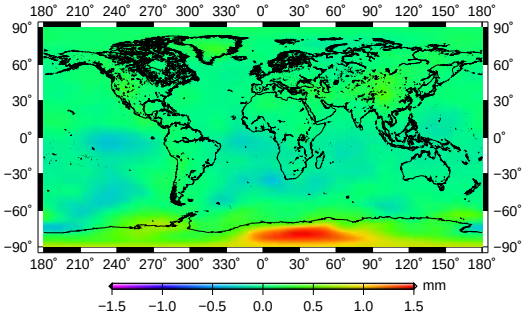
(b) Mean ECMWF surface pressure, IB, in April, 2001



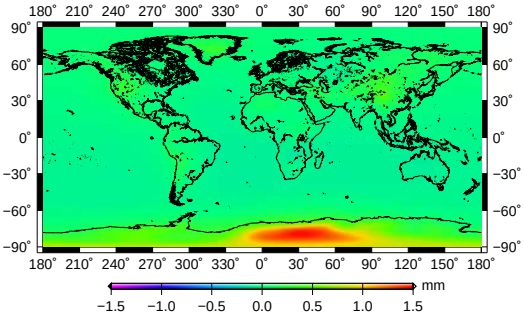
(c) Mean NCEP reanalysis surface pressure, non-IB, in April, 2001



(d) Mean NCEP reanalysis surface pressure, IB, in April, 2001

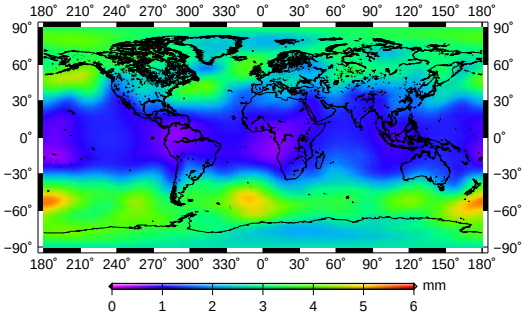


(e) Mean difference between ECMWF and NCEP reanalysis, non-IB, in April, 2003

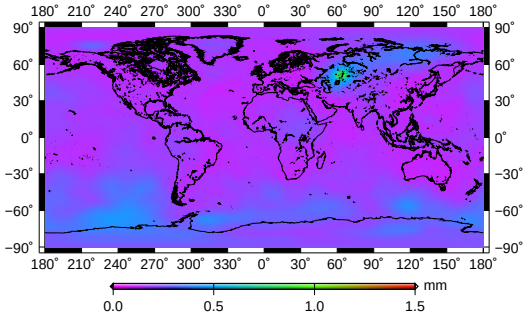


(f) Mean difference between ECMWF and NCEP reanalysis, IB, in April, 2001

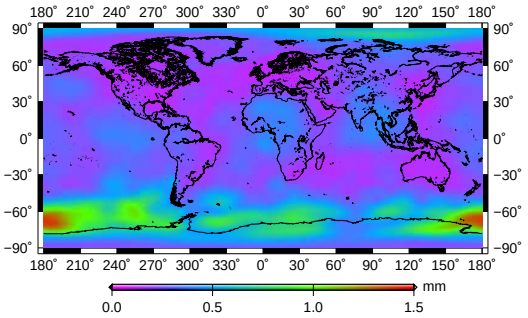
Figure 5.12: The monthly mean surface pressures in the simulation.



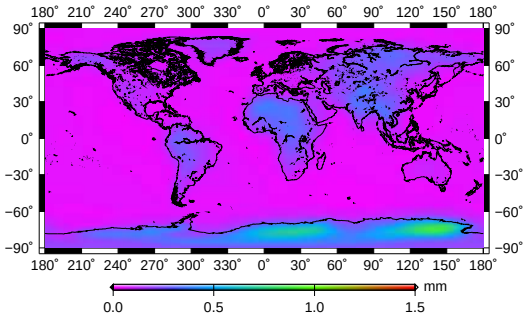
(a) RMS ECMWF surface pressure, non-IB, in April, 2001



(b) RMS of the difference between ECMWF and NCEP operational analysis, non-IB, in April, 2003



(c) RMS of the difference between ECMWF and NCEP reanalysis surface pressure, non-IB, in April, 2001



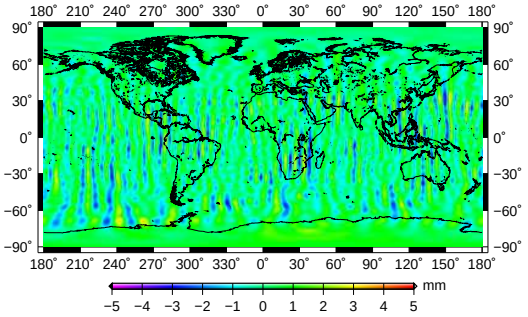
(d) RMS of the difference between ECMWF and NCEP reanalysis surface pressure, IB, in April, 2001

Figure 5.13: The monthly RMS surface pressures in the simulation.

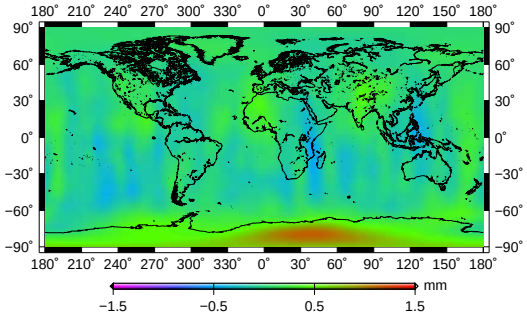
5.15) shows that the aliasing errors spread over the globe, which indicates that the variations in one area could introduce aliasing error in other areas. Improving the pressure modeling in Antarctica is therefore helpful to reduce the global aliasing error.

5.6.5.4 Summary and discussion

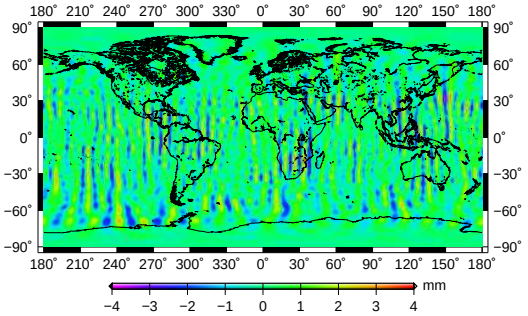
We investigate the atmospheric aliasing error using ECMWF TOGA analysis as the true model, and NCEP reanalysis and NCEP operational analysis as de-aliasing models. Both IB and non-IB are considered. It is found that there are more noticeable improvements for the middle to high degree coefficients after applying a de-aliasing model. From the simulation result, most of the long wavelength errors can be correctly captured by GRACE as mis-modeled or un-modeled errors. The motion of the long wavelength features induces short period, small scale variations, which will further



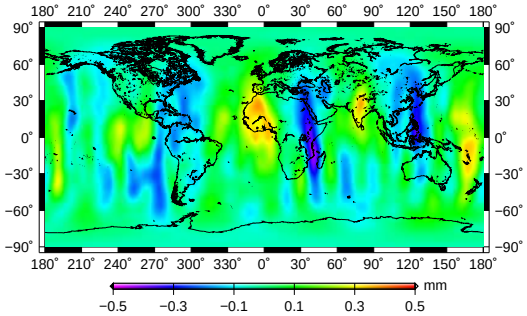
(a) GRACE recovered ECMWF-NCEP reanalysis+noise in April, 2001



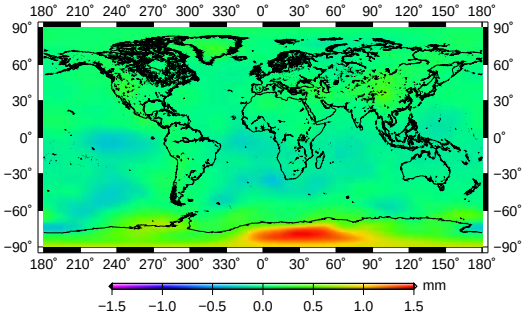
(b) 600 km smoothing of (a)



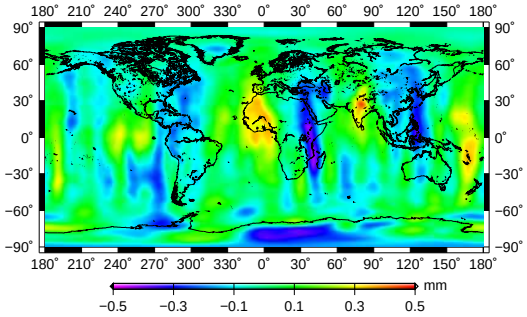
(c) Aliasing error (a)-(e)



(d) 600 km smoothing of (c)



(e) True surface pressure error, monthly mean ECMWF-NCEP in April, 2001



(f) Difference between (b) and (e)

Figure 5.14: The simulation results for the atmospheric de-aliasing.

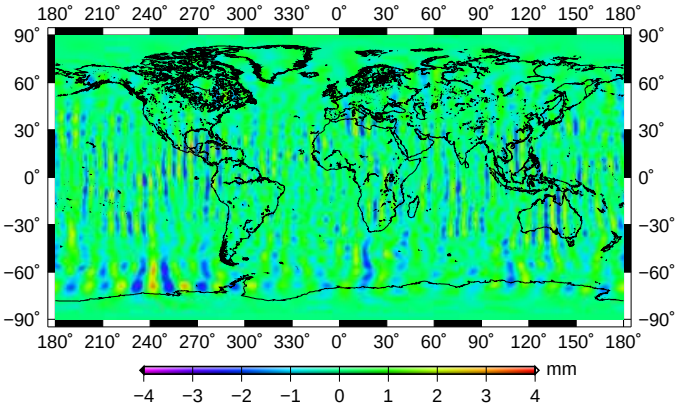


Figure 5.15: The aliasing error caused by the pressure variations in Antarctica only.

increase the errors for the higher degree terms. The magnitude of the aliasing error is correlated with the RMS of the pressure field. The simulation using the NCEP operational analysis in 2003 shows relatively smaller aliasing error. It implies that the NCEP model agrees better with the ECMWF than before and the model is continuously improving.

Under the IB assumption, the degree variance of the pressure field becomes much smaller compared to the non-IB case, and so does the geoid change. Large variations around the Southern Ocean area are neglected, which, in turn, leads to smaller aliasing error. By neglecting the pressure over the ocean, the global mean value ($C_{0,0}$) is changed at each time. Since the degree 0 and 1 terms are neglected, the RMS value over land areas under IB assumption is slightly different from the land RMS value under non-IB assumption. The aliasing error caused by the pressure variations over the land is different in these two cases. Due to the truncation of the spherical harmonics, a strong Gibbs phenomenon is present around the continents. Some variations can leak into the oceans. The IB assumption might lead to an optimistic estimate of the aliasing error, while the non-IB assumption can give us a more conservative error estimation.

Spatial smoothing greatly reduces the amplitudes of both the un-modeled/mismodeled and aliasing errors with the sacrifice of the amplitude of the signal. The aliasing is not a localized phenomenon. Variations in one area might relate to the aliasing error in a remote area.

The uncertainty of the surface pressure measurements in the Southern Ocean and Antarctica is very big. The best way to obtain precise surface pressure in data-poor regions is to extend the current barometric network. Because of the expense for the instruments and their maintenance, this suggestion is less favorable and impractical. GPS occultation, as a promising technique, provides a potential to improve the

pressure field over the Southern Ocean and Antarctica. It already appears in many publications that the accuracy of the GPS-occultation derived profiles is commensurate with the analysis model output and radiosonde measurements. We also believe that surface pressure derived from GPS occultation can provide information with the same accuracy. GPS occultation can be treated as another independent measurement to validate other sources of surface pressure measurements. Another application is to directly combine the occultation measurements with a numerical analysis model, i.e. to assimilate the bending angle or refractivity into the numerical models. Some studies, e.g., [*Zou et al.*, 1999], show that the assimilation of atmospheric refractivity is very effective in recovering the vertical profiles of water vapor. The accuracy of the derived water vapor field is significantly better than that obtained through traditional retrieval techniques. This would, in turn, improve the surface pressure fields.

CHAPTER 6

THE EVALUATIONS OF NWP MODELS IN ANTARCTICA USING AUTOMATIC WEATHER STATION DATA

6.1 Introduction

For higher southern latitudes, especially the Antarctic continent, the numerical model outputs are not as consistent as in other parts of the world. Our comparisons of pressure in the southern polar region show large discrepancies for different models. The lack of observation data is counted as a major obstacle for the improvement of model performance. In Antarctica, the extreme weather conditions and special geographical location present big challenges for standard meteorological observations. Since the area is less accessible, the calibration and maintenance of instruments sometimes need extensive periods of time. Instrument failure and communication problems further reduce the quality and quantity of the data. In addition, due to the sharp topography change, which rises quickly inland, and extremely cold weather over the ice sheet, the physical modeling and numerical representation are quite different from the situation in the mid-latitudes. These factors jointly degrade the model performance.

For the GRACE mission, the mass balance signal is contaminated by the atmospheric pressure residual error. Theoretically, 1 hPa pressure error causes a mass change represented by water thickness of 1 cm. The ECMWF pressure is adopted to remove the atmospheric mass change. *Trenberth* [1992] evaluated the ECMWF and NCEP analyses during 1979-1986, and found major problems in both datasets in the southern polar region. *Cullather et al.* [1997] compared the ECMWF and NCEP analyses with observations during the period 1985-1994, and it is found that ECMWF pressure correlates reasonably well with the AWS data, and NCEP analyses show substantial improvement. *Velicogna et al.* [2001] validated the surface pressure over the U.S. continent and Arabic peninsula, and the authors concluded that the model output from ECMWF is adequate for the GRACE mission over such regions.

It is also pointed out in the paper that model outputs have problems in the southern polar region, but without further validation. New types of satellite observations are continually incorporated, and the operational models are routinely evolving with updated and improved methods. However, trends could be introduced with the alteration of data and model algorithm [e.g. *Marshall*, 2002; *Hines et al.*, 2000]. Therefore, the continuous monitoring and assessment of the model pressure during the GRACE period is necessary. In this chapter, the accuracy of the surface pressure from the most recent ECMWF and NCEP re-analysis is validated against the observations from the automatic weather stations.

6.2 The Antarctica weather

Most of the Antarctic continent is situated within the Antarctic Circle (60°S of the Earth). The continent has an area of 1.4×10^7 km^2, consisting of nearly 10% of the Earth's total land area. Nearly 97% of the area is covered with ice, which greatly affects the heat balance of the continents. Most of the solar (short wave) radiation energy is reflected back because of the high albedo of the ice. The long wave energy is further lost due to the high emissivity of the ice. The annual average temperature is around -35°C. The temperature is only above freezing in the west side of the Antarctic Peninsula. The low surface temperature causes strong temperature inversion in the boundary layer.

The weather in Antarctica is greatly affected by its topography, which has a notably asymmetric feature. The average elevation is around 2,000 m, with most of the west Antarctica below 2,000 m except some mountain regions with elevation of 3,000 m. The highest elevation is on the east Antarctica ice sheet, with a central dome over 4,000 m. Another feature of Antarctica is that the elevation rises rapidly from the coast. It could reach 2,000 m just within 500 m distance. The topography has a strong influence on the pattern of temperature, surface wind, pressure, and precipitation. The correlation between the elevation and temperature is very strong. In the coast and peninsula area, the temperature has a regular pattern of a maximum in summer and a minimum in July and August. Further inland, a maximum summer season exists, but with a "coreless" winter with little temperature variation. The strong surface cooling at the boundary layer causes the dense, cold, and dry air flow downhill following the slope of the topography. This gravity induced air flow (katabatic wind) is the dominant wind pattern, which is more intense in the winter season. The elevation also forms a natural barrier and prevents most of the synoptic weather system from penetrating into the interior of Antarctica. Most of the heavy precipitation is concentrated along the coast line, where the adiabatic cooling as the moist air from ocean climbs up the steep topography is the major mechanism for the formation of precipitation.

The dominant feature of the Mean Sea Level Pressure (MSLP) in the Antarctica region is the circumpolar trough, which is usually situated between the latitude 60°S – 70°S. The circumpolar trough oscillates semi-annually in both position and time [*Simmonds and Jones*, 1998]. The trough is the deepest in spring and autumn and the most shallow in summer and winter. However, the reduction of the surface pressure to mean sea level is less accurate and less reasonable due to the high elevation and the cold environment. 500 hPa is the lowest standard pressure level that is everywhere above the Antarctic continental surface. Generally, the pressure is higher when moving toward the interior. Surface pressure can decrease up to 20 hPa over the continent during a four day period [*Parish and Bromwich*, 1998]. The seasonal variations of surface pressure are more obvious in the coast than that in the interior. Figure 6.1 shows the monthly mean surface pressure of AWS in Antarctica for 2003, which are grouped according to their geographical regions. A mean value is removed from each time series. In the figure, the maxima are clearly shown around the summer and winter solstices. It indicates that there is a net transport of air into high latitude in the early winter and spring, and corresponding export of atmospheric mass to lower latitudes in late winter and summer [*King and Turner*, 1997].

6.3 The automatic weather station data

The extreme cold, icy environment in Antarctica posts a huge challenge for instruments to work in normal situations. Sometimes, the wind and snow can bury, even damage the instruments. It is very expensive to use man-operated stations to obtain *in situ* measurements. Most of the manned stations are located in the external margins of the Antarctic coast line. There are few stations in the interior of the Antarctic continent. The unmanned AWS makes a large supplement to the high-cost research stations, and provides valuable surface observations to support weather research. Since 1980, a network has formed with a growing number of stations. The number of stations is not stable due to the ending of various projects and equipment failures. Currently, there are under 60 stations in operation, which are clustered in some research areas. Most of them are along the coast. There are only a few stations in East Antarctica. Typically, the AWS units measure wind speed, direction, and air temperature at a nominal height of 3 m above the surface. The air pressure is measured at an electric enclosure at the height of about 1.75 m. The height varies from time to time due to snow accumulations. The equipment is powered by battery and solar panels. It can usually work through to next field season. It is reported that some stations can work on the same batteries and solar panel for 6-10 years. The stored data are transmitted to two NOAA satellites with an orbital period of 102 minutes. The accuracy of the pressure transducer is around 0.25 hPa [*King and Turner*, 1997]. Figure 6.2 shows a schematic figure for a typical AWS.

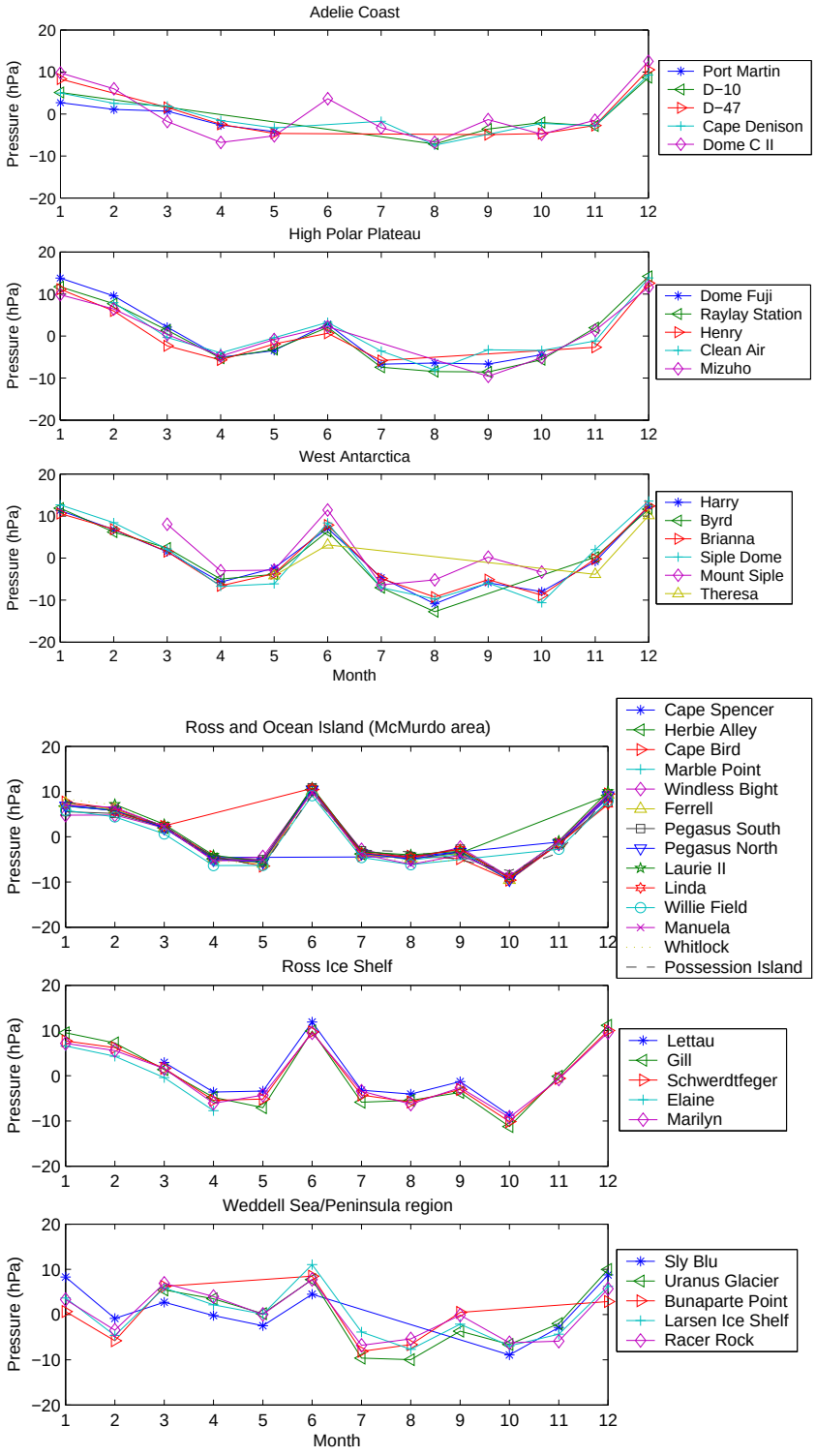


Figure 6.1: Monthly mean AWS surface pressure in different regions of Antarctica (Year 2003). A yearly mean was subtracted from the monthly mean for each time series.

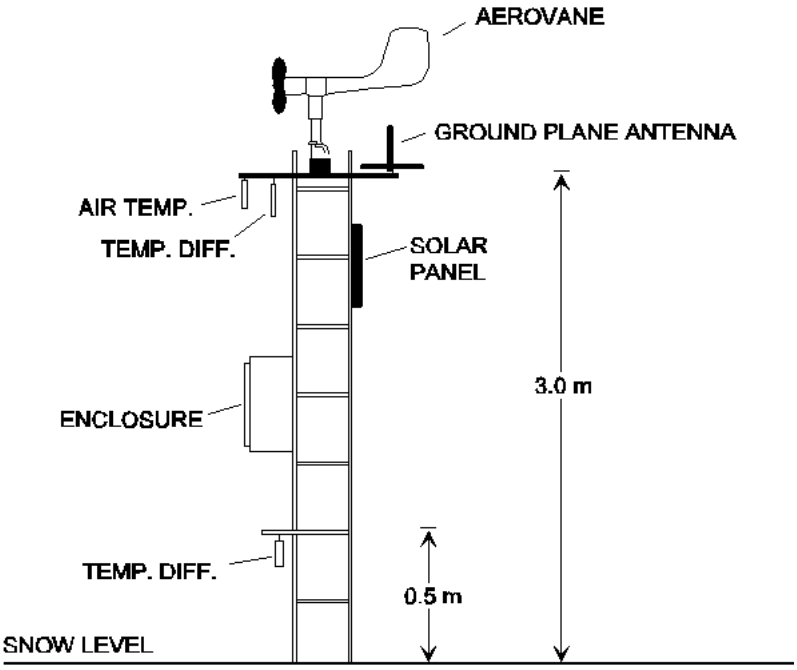


Figure 6.2: A schematic figure for Automatic Weather Stations, figure is from the Wisconsin AWS project.

The automatic weather stations project at the Antarctica Meteorological Research Center of The University of Wisconsin-Madison maintains an archive of AWS data. The data are available from its anonymous ftp site. The AWS 10-minute data are available in raw format without being corrected for bad data. There are also 3-hour data available after some extensive period, which has been corrected for bad observations by hand. However, after 1998, the 3-hour data are not readily available. As an alternative, a whole year of 10-minute data for 2003 is downloaded. The "raw" data have various data gaps in it; and sometimes there are numbers of unreasonable data records. The missing and invalid records are probably mainly caused by communication problems. Manually fixing the data would take a lot of work. A program is designed by first filling the gap using linear interpolation if the gap is less than 6 hours. Our model data is available every 6 hours. If the data gap is longer than 6 hours, the time series for the whole year is divided into segments. For each segment, the outliers are detected, removed and fixed by a polynomial using the singular value decomposition method. At the final step, some outliers cannot be identified or removed. These remaining bad data are removed by manual editing. When comparing AWS with the model data, a 6-hour time series is excerpted from the corrected 10-minute AWS data.

6.4 Temporal interpolation of analysis results

Most of the meteorological analysis models only produce output every 6 hours. The processing of GRACE data needs to remove pressure contribution every couple of seconds. A temporal interpolation is required within the 6-hour period. We need to examine how much information is included in the 6-hour data and whether a simple linear interpolation can fulfill the requirement of the accuracy.

We randomly select a station and plot the time series of surface pressure of 10-minute AWS, 6-hour ECMWF, and NCEP in Figure 6.3. The figure shows that the analysis results follow the variation of the surface pressure very well, though they do not agree exactly. Notable differences occur at some peaks and valleys of the curve.

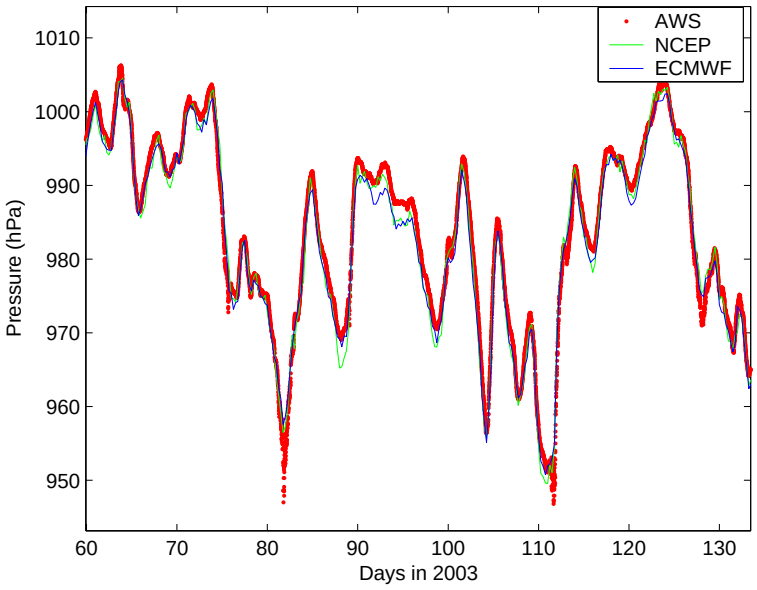


Figure 6.3: AWS, ECMWF, and NCEP Surface pressure records at Cape Bird.

We also examine the Power Spectrum Densities (PSD) of AWS and ECMWF datasets with distinct sampling intervals. Figure 6.4 shows the PSDs of 10-minute AWS data and 6-hour ECMWF output for two selected AWS locations. These two AWS stations have the smallest missing data in the one year period, and the gaps are filled using linear interpolation. The Nyquist frequency of 6-hour ECMWF data is marked by a vertical red line at 2 cycles/day. Below the Nyquist frequency, the PSD curves agree very well. The tail of PSD after the Nyquist frequency is not included in the ECMWF data. The missing power is the integration of this tail. For these two particular stations, the missing power is on the order of 0.01 hPa, which corresponds

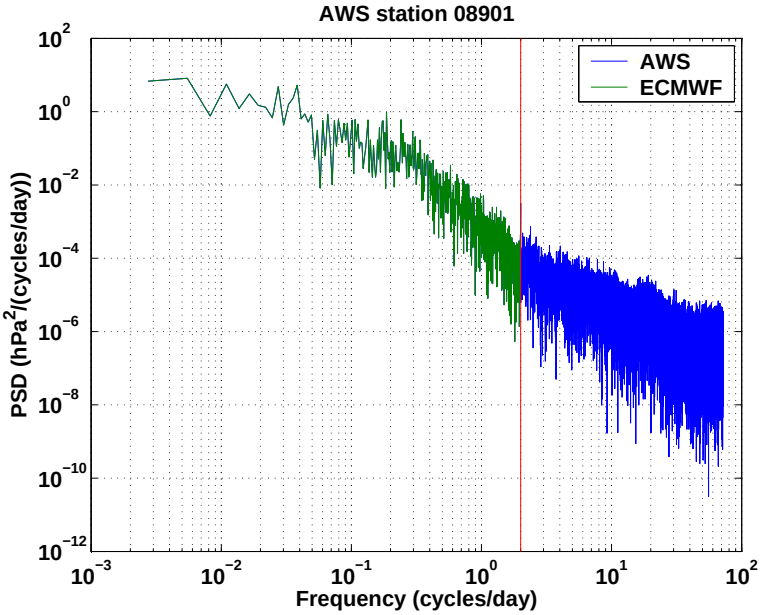
to 0.1 mm equivalent water thickness. It is negligible for current GRACE sensitivity. We conclude that 6-hour data can capture most of the pressure variations, and a simple linear interpolation is sufficient for higher sampling data.

6.5 Comparisons of surface pressure data

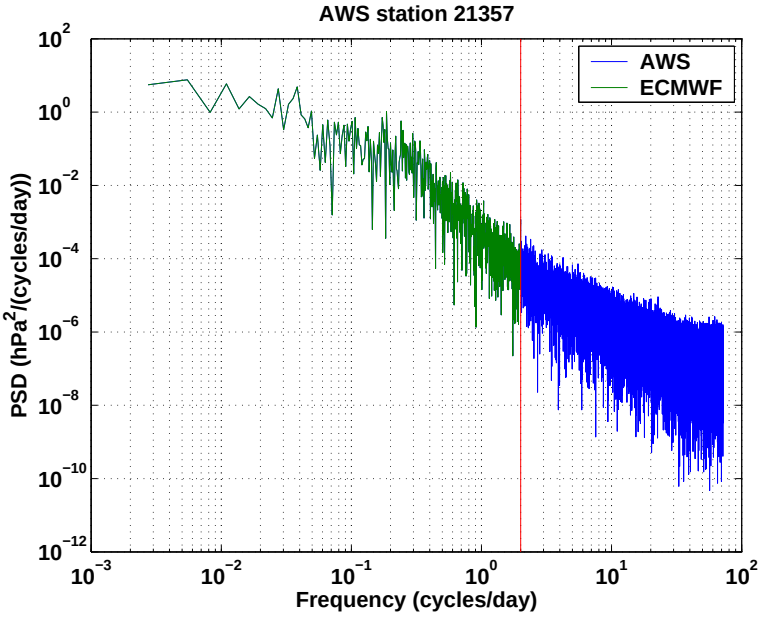
As stated above, we use the corrected AWS surface pressure series for this comparison. The data from 40 stations are included in this study. The geographic locations for these stations are shown in Figure 6.5. The detailed information for each station is listed in Appendix C. For the analysis datasets, ECMWF TOGA global advance surface analysis from the NCAR data center and NCEP reanalysis data from the NOAA Climate Diagnostics Center are used. The ECMWF dataset is in N80 Gaussian grid and stored in GRIB format, with 1.125° in the longitude direction and approximately 1.125° in the latitude direction. This dataset is interpolated later to a 1°×1° regular grid. The NCEP reanalysis data is on a 2.5°×2.5° regular grid and stored in netCDF format.

Directly comparing the model surface pressure with the AWS surface pressure is questionable since the model elevation is usually not the representation of the real Earth surface. In the coastal area, the difference is rather big. *Bromwich et al.* [1999] reported that the difference could be as large as 140 hPa in the steeply topographic area. In the last two columns of Table 6.1, the topographic differences between ECMWF and AWS as well as NCEP and AWS are listed. For some stations, the difference is as high as 900 m. NCEP reanalysis data has a coarser resolution, and the corresponding difference is larger. In the past, the AWS elevations were determined by a variety of inaccurate methods, that were best at the time, including using the older GPS system onboard an aircraft that was deployed to the station locations, and using USGS elevation maps. Recently, there have been efforts to tie the AWS stations to the UNAVCO GPS locations in Antarctica. This is an ongoing process and can only be applied to the stations that can be visited. In Appendix C, the AWS stations used in this study are listed. Sites that have been surveyed by UNAVCO are marked in the table.

The monthly mean difference between AWS and ECMWF is shown in Figure 6.6. The bias between AWS observations and ECMWF is almost constant across the year, which implies that there are some systematic errors between the two datasets. The most possible cause is the elevation error, either in the model or in the nominal height of AWS. When the elevation rises approximately 10 m near the ground, the pressure changes 1 hPa. There are better agreements for the stations in the high polar plateau and Ross Ice Shelf. On the contrary, the biases for some stations in the peninsula area are rather big ($>$ 20 hPa). If a mean is removed from each time series before taking the difference, the two series agree better. The bias is not an exact constant value, it does vary in time as depicted in Figure 6.7a at station Dome C II. To determine



(a) Cape Bird



(b) Pegasus North

Figure 6.4: The power spectrum densities of AWS surface pressure observations (10 minute interval) and ECMWF surface pressure at AWS locations (6 hour interval) for year 2003.

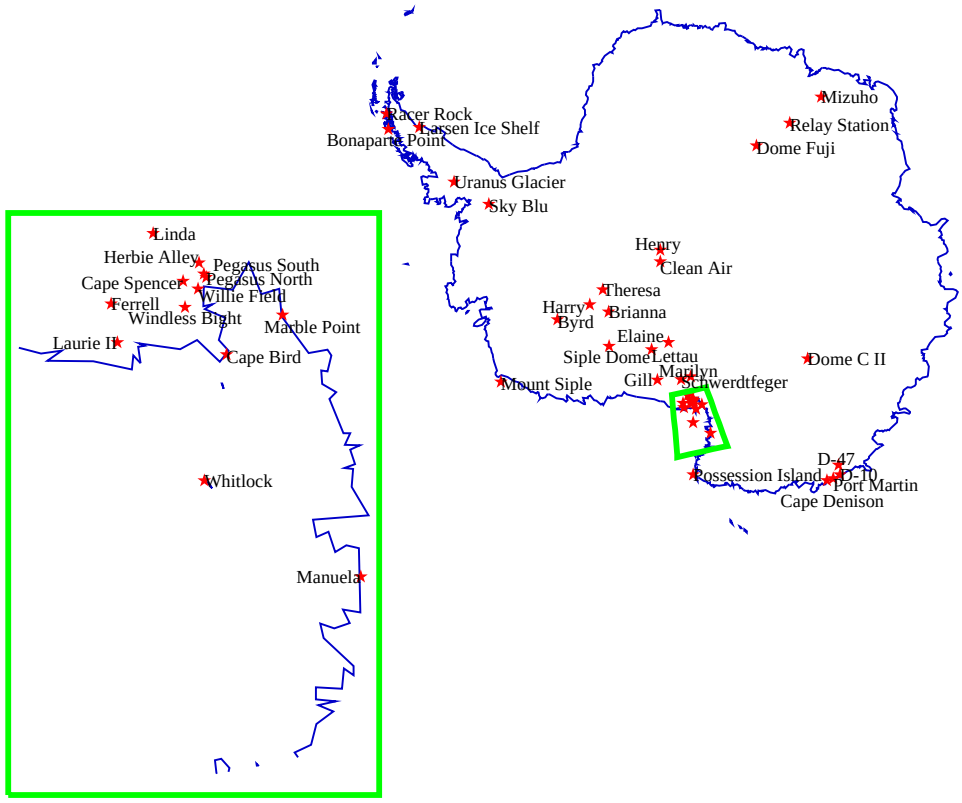


Figure 6.5: The locations of 40 AWS stations.

whether the bias has periodicity we need to analyze at least two years of data. For the monthly mean difference between AWS and NCEP, Figure 6.8 shows similar patterns, though the bias curves are not as flat as that in the ECMWF case. The biases in NCEP have the similar sign and amount of magnitude as in ECMWF, which implies that the two analysis models are also correlated.

The comparison of the mean difference is hindered by the uncertainty of the AWS elevation height. The standard deviation of these monthly mean differences, on the other hand, may reveal the real agreement between the observations and models. In Figure 6.9, some stations at coastal regions have relatively higher standard deviation values. Stations at the high polar plateau, west Antarctica and Ross Ice Shelf have standard deviation values less than 2 hPa. We examine the stations that have higher standard deviation values and find that the AWS records have large fluctuations around those times (see Figure 6.7b). From the figures, the big values are usually around winter time. Possibly some weather activities are developed. The standard deviation comparison for the monthly difference between NCEP and AWS shows a different picture (Figure 6.10). There is no similarity with the ECMWF comparison.

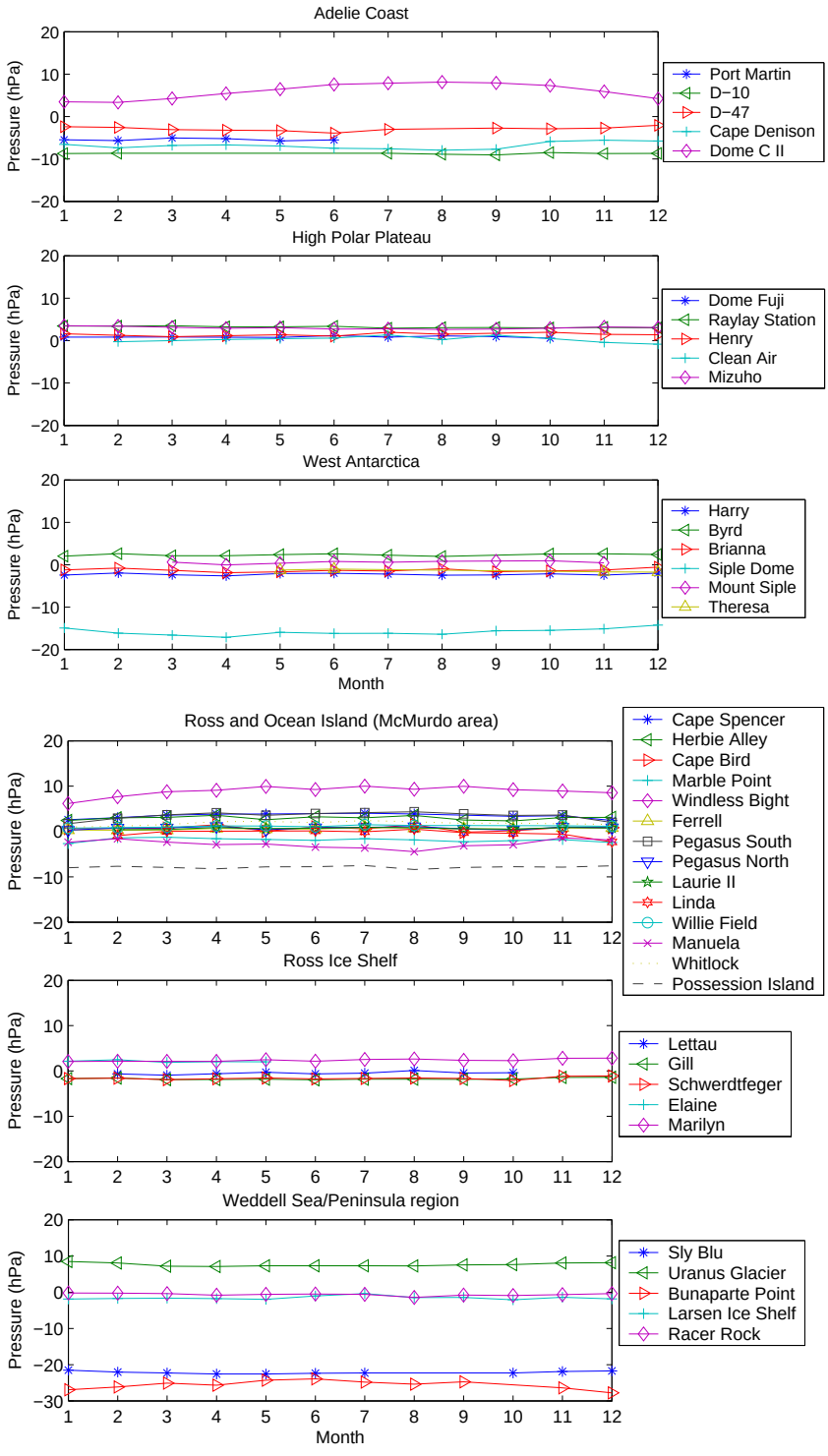
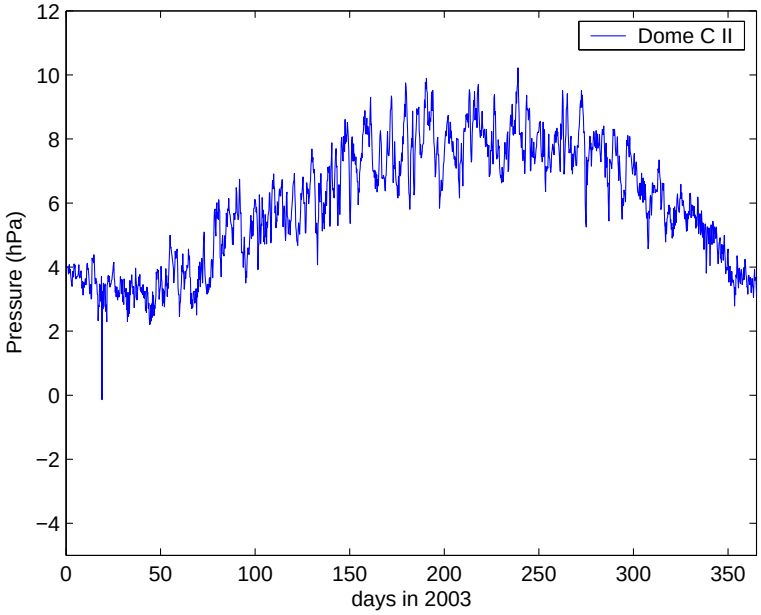
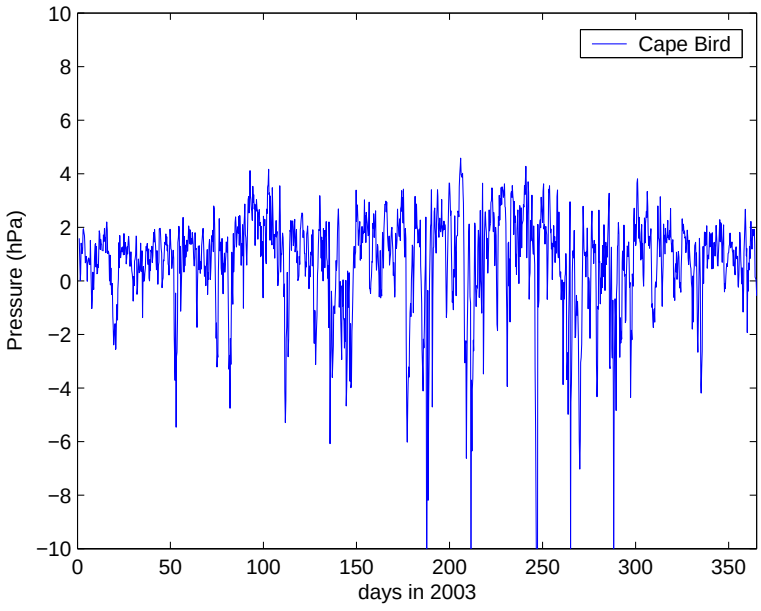


Figure 6.6: Monthly mean of the surface pressure difference between AWS and ECMWF at AWS stations (Year 2003).



(a) AWS-ECMWF pressure difference at Dome C II



(b) AWS-ECMWF pressure difference at Cape Bird

Figure 6.7: The surface pressure differnce between AWS and ECMWF at two AWS stations.

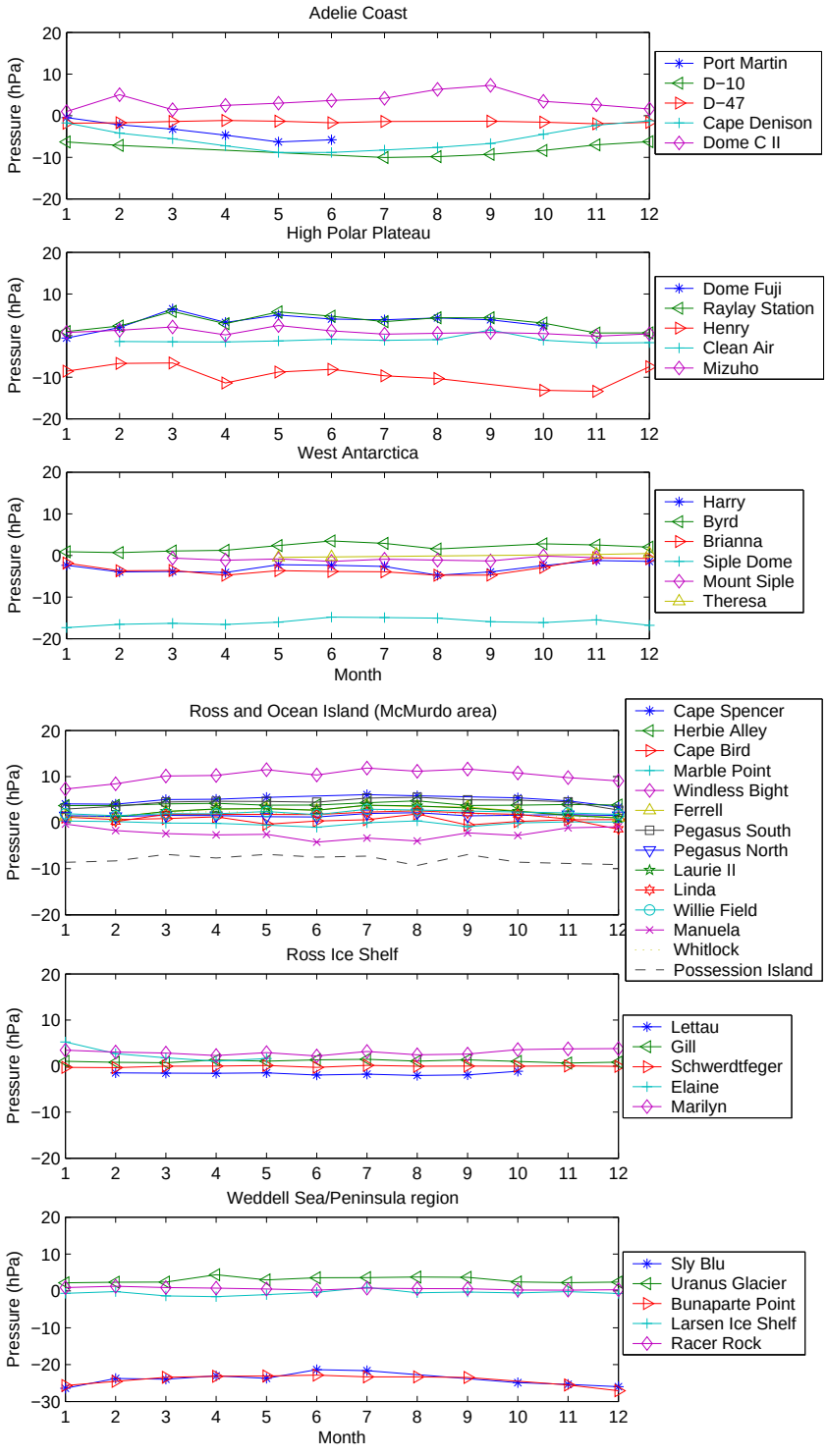


Figure 6.8: Monthly mean of the surface pressure difference between the AWS and NCEP re-analysis at AWS stations (Year 2003).

The values are large even in the interior region. There is almost no seasonal difference. It is clear that the agreement between ECMWF and AWS is much better than that between NCEP and AWS.

In a past study [*Ge et al.*, 2005], we also compared one month of surface pressure from Polar MM5 [*Bromwich et al.*, 2001] with the ECMWF and the surface AWS measurements. The MM5 surface pressure is available in a 60 km stereographic projection grid and has a temporal resolution of 1 hour. The spin-up time for the model is 24 hours. The output from MM5 is 24-48Z (Z stands for Greenwich time) forecast, which achieves a reasonable performance, but the difference is still large at some time. Since MM5 outputs are prediction results and there is no other input except the initialization, its agreement with AWS is not better than that of the ECMWF.

Generally, the statistical patterns should be similar within same region. There are also some times, when the meteorological sensor reports suspicious values and causes large standard deviations. In addition, these values vary from month to month, indicting a possible correlation with seasonal synoptic weather activities. The standard deviation for the whole year of the difference between ECMWF and AWS is plotted on a geographic map in Figure 6.11. The numerical values are listed in the 4th column of Table 6.1.

6.6 Summary

We compare the mean and standard deviation of the difference between model analyses and AWS surface pressure observations using one year of data. ECMWF shows a much better agreement with the AWS in terms of standard deviation. The uncertainty in elevation causes large bias between the model and AWS. The bias is not always a constant throughout the year. Part of the bias can be removed by subtracting a yearly mean. The statistics reveal the spatial and temporal distributions of errors. Coastal regions tend to have higher errors than the interior areas, which correlated with the Antarctica weather pattern, i.e. the synoptic weather in coastal areas is more active. Another reason is that the spherical model poorly represents the steep change of the topography. Large disagreements also tend to happen in winter and in regions with increasing weather activities. In Antarctica, the standard deviation of the pressure difference is around 2 hPa in the interior and larger than 2 hPa along the coast. One should also note that AWS data may already have been assimilated into the models and are not fully independent. This makes our analysis too optimistic. In this study, other factors such as different model resolutions also affect the comparison results.

In the GRACE data processing, a mean value of at least one year is subtracted from the 6-hour field, which alleviates part of the bias problem if the bias is a constant. In the case where the difference varies with time, the error source has to be identified.

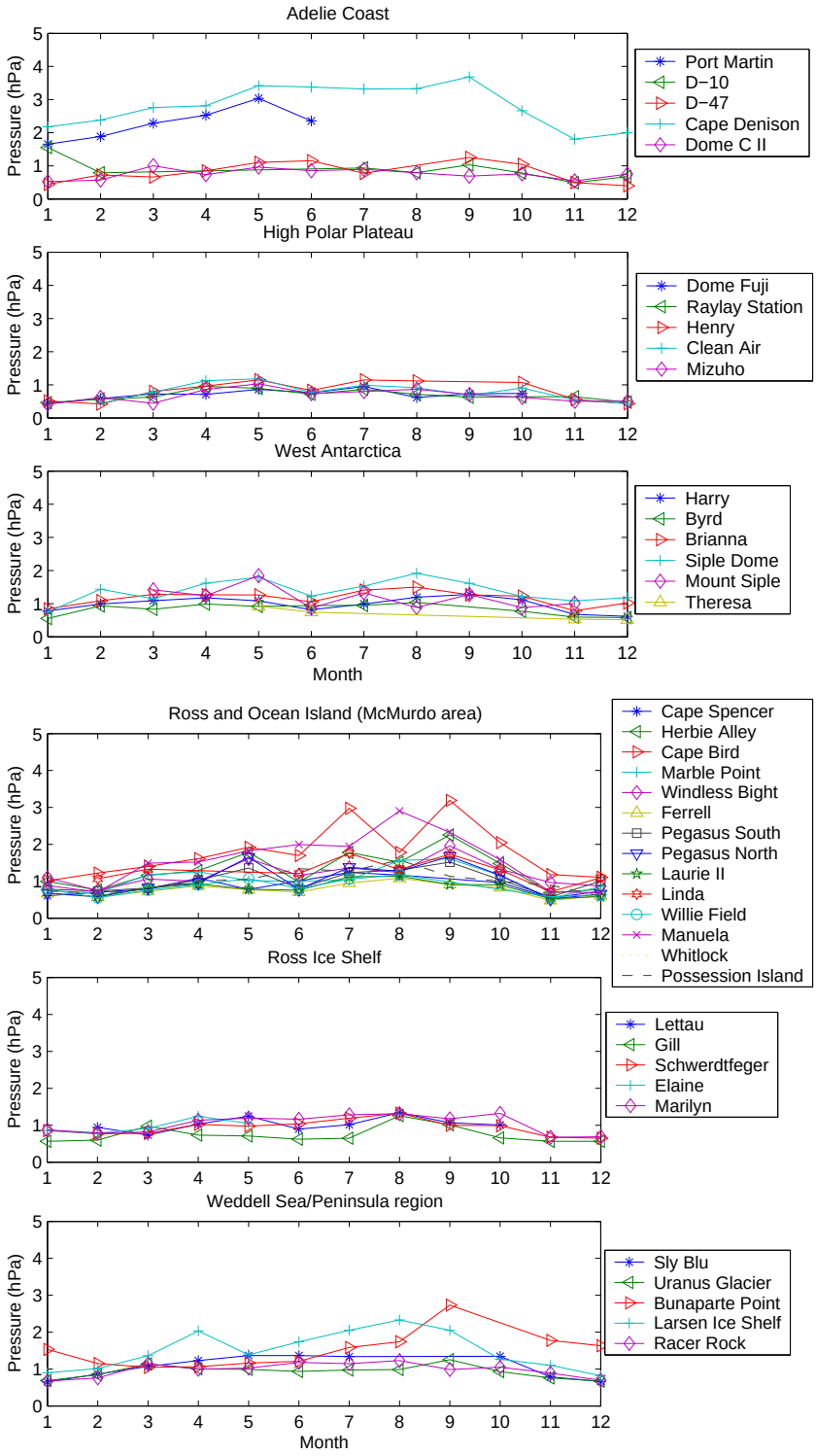


Figure 6.9: Monthly standard deviation of the surface pressure difference between the AWS and ECMWF at AWS stations (Year 2003).

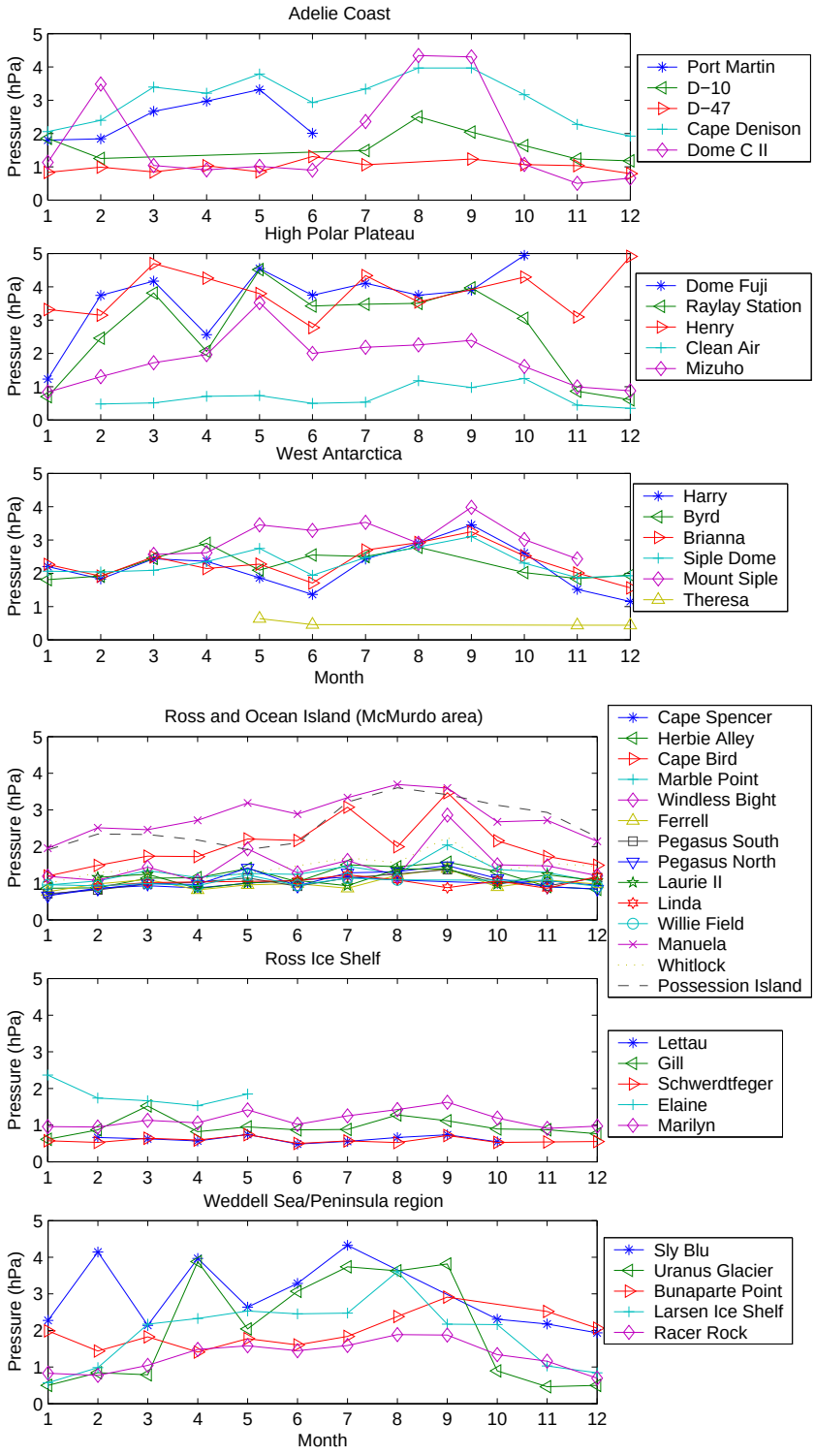


Figure 6.10: Monthly standard deviation of the surface pressure difference between the AWS and NCEP re-analysis at the AWS stations (Year 2003).

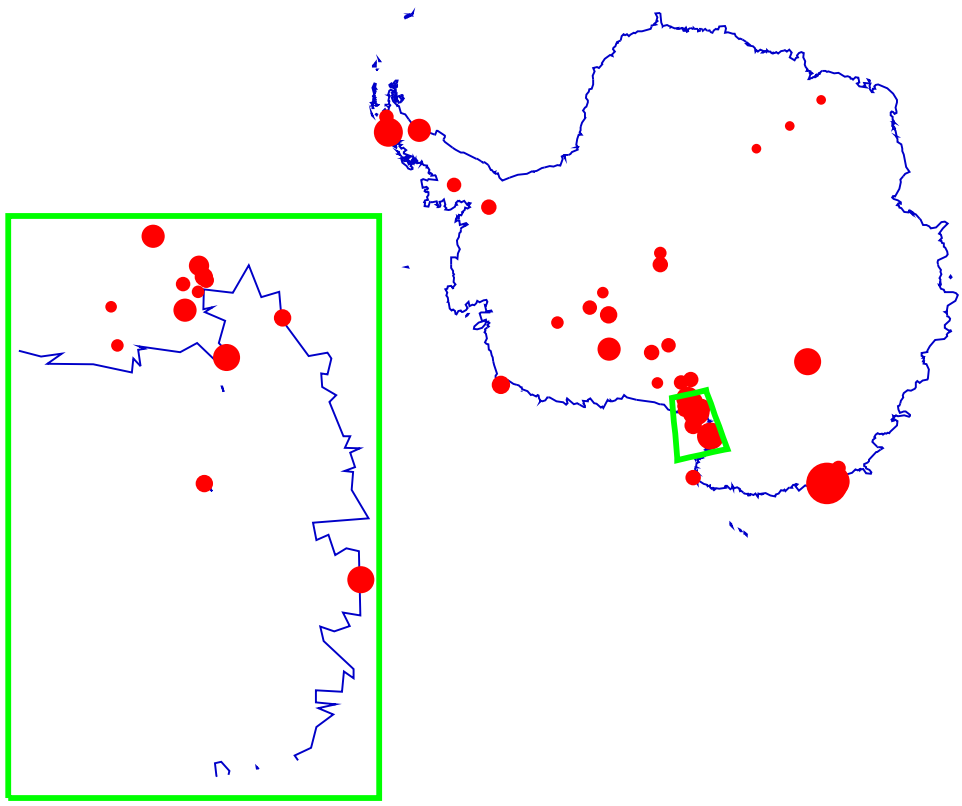


Figure 6.11: The yearly standard deviation for the surface pressure difference between ECMWF and AWS.

The above statement is based on the assumption that the process is linear. For a general gravity inversion problem, there are many non-linear processes. Further simulations are needed to quantify the errors. From our comparisons, the error is not as good as in other parts of the world. Due to the limitation of the data and clustering of the stations, we cannot calculate a large area mean (e.g. Gaussian smoothing) in terms of a radius as regularly used in GRACE. A larger than average error budget has to be expected in Antarctica.

Since there is usually no accuracy information given for the analysis models, people usually use the difference between ECMWF and NCEP as an indicator of errors. These two fields always have something in common, such as model algorithm, data, etc. The common errors are canceled by subtracting one from the other, and thus, the difference between these two models may not represent the true error distribution. Topographic difference is one such contribution. In previous chapters, large surface pressure differences are seen in the interior part of Antarctica, which does not agree

with the difference between ECMWF with AWS. If we can trust the AWS observations, the statistics of the difference could be used as a valuable error indicator.

Station Name	obs. num.	AWS-ECMWF ave	std	corr.	AWS-NCEP ave	std	corr.	$h_{ecm}-h_{aws}$	$h_{ncp}-h_{aws}$
Adelie Coast									
Port Martin	658	-5.4	2.3	0.9612	-3.6	3.3	0.9231	659.43	615.38
D-10	931	-8.7	0.9	0.9962	-8.0	2.2	0.9789	374.58	502.53
D-47	1320	-2.9	1.0	0.9959	-1.5	1.0	0.9943	-269.19	-412.27
Cape Denison	1385	-6.8	2.9	0.9632	-5.5	4.1	0.9274	579.64	572.71
Dome C II	1455	6.0	1.9	0.9826	3.5	2.9	0.9514	-37.11	-170.23
High Polar Plateau									
Dome Fuji	1256	0.9	0.7	0.9973	3.2	4.2	0.9280	-77.96	-638.00
Relay Station	1455	3.2	0.7	0.9975	3.2	3.5	0.9506	-62.44	-475.27
Henry	1286	1.4	0.9	0.9955	-9.4	4.6	0.8791	-60.67	-914.17
Clean Air	1331	0.3	1.1	0.9934	-1.1	1.1	0.9925	13.54	68.00
Mizuho	1383	3.0	0.7	0.9972	0.8	2.1	0.9756	-28.80	-383.40
West Antarctica									
Harry	1456	-2.3	1.0	0.9957	-2.9	2.5	0.9738	11.88	224.57
Byrd	1295	2.3	0.9	0.9970	1.9	2.5	0.9757	-14.06	-77.74
Brianna	1455	-1.3	1.2	0.9945	-3.2	2.7	0.9723	14.41	228.09
Siple Dome	1456	-15.8	1.6	0.9919	-16.0	2.5	0.9810	-412.99	-505.39
Mount Siple	1080	0.6	1.3	0.9959	-0.9	3.1	0.9723	-38.34	-30.62
Theresa	601	-1.5	0.8	0.9963	-0.1	0.7	0.9976	260.68	217.33
Ross Ice Shelf									
Lettau	1138	-0.5	1.1	0.9962	-1.6	0.7	0.9986	-29.00	-63.71
Gill	1456	-1.7	0.8	0.9980	1.1	1.0	0.9968	-25.03	-193.11
Schwerdtfeger	1456	-1.6	1.0	0.9965	-0.1	0.6	0.9987	-56.99	-32.73
Elaine	541	2.1	1.0	0.9964	2.6	2.4	0.9788	169.99	590.15
Marilyn	1456	2.4	1.1	0.9952	3.0	1.3	0.9935	-70.46	364.97
Weddell Sea/Peninsula Region									
Sky Blu	1222	-22.2	1.1	0.9936	-24.1	3.4	0.9468	-63.23	-666.84
Uranus Glacier	1376	7.6	1.0	0.9970	3.1	2.6	0.9792	-255.70	7.95
Bonaparte Point	1242	-25.5	2.0	0.9886	-24.2	2.4	0.9827	157.89	160.04
Larsen Ice Shelf	1454	-1.6	1.6	0.9900	-0.5	2.2	0.9840	-19.39	162.85
Racer Rock	1452	-0.6	1.0	0.9968	0.6	1.4	0.9943	147.49	146.16

Continued

Table 6.1: The statistics of automatic weather stations in Antarctica for 2003

Table 6.1 Continued

Station Name	obs. num.	AWS-ECMWF ave	AWS-ECMWF std	AWS-ECMWF corr.	AWS-NCEP ave	AWS-NCEP std	AWS-NCEP corr.	$h_{ecm}-h_{aws}$	$h_{ncp}-h_{aws}$
Ross and Ocean Island (McMurdo Area)									
Cape Spencer	1124	3.4	1.0	0.9956	4.9	1.3	0.9941	37.01	153.86
Herbie Alley	1424	3.0	1.4	0.9921	4.0	1.2	0.9941	14.91	251.80
Cape Bird	1456	0.7	1.9	0.9857	0.6	2.2	0.9810	-31.53	275.94
Marble Point	1456	-1.9	1.2	0.9943	-0.1	1.4	0.9921	343.49	547.37
Windless Bight	1453	8.9	1.6	0.9892	10.2	2.0	0.9849	-10.22	99.86
Ferrell	1456	0.6	0.8	0.9974	2.5	1.4	0.9940	-51.64	-103.36
Pegasus South	1456	3.5	1.3	0.9926	4.4	1.3	0.9935	37.11	281.06
Pegasus North	1456	0.7	1.1	0.9950	1.6	1.1	0.9952	31.63	285.67
Laurie II	1331	0.7	0.9	0.9971	2.6	1.4	0.9933	-42.17	-103.38
Linda	1166	-0.4	1.6	0.9894	1.3	1.6	0.9900	-8.43	75.53
Willie Field	1232	1.1	0.9	0.9966	2.1	1.2	0.9950	27.73	224.06
Manuela	1456	-2.7	1.9	0.9858	-2.4	3.1	0.9604	233.98	835.50
Whitlock	1456	1.8	1.2	0.9945	1.9	1.6	0.9904	-272.51	-100.40
Possession Island	1453	-7.9	1.1	0.9948	-8.0	2.8	0.9661	-35.55	245.50

CHAPTER 7

MATHEMATICAL DESCRIPTION OF GRACE ATMOSPHERIC DE-ALIASING MODEL AND ITS VALIDATION

The GRACE mission requires the accurate removal of the atmospheric mass variations. The contributions of atmospheric mass to the geopotential include the direct mass attraction and the indirect loading effect caused by the deformation of the Earth. Conventionally, the atmosphere is treated as a dimensionless thin layer around the spherical Earth. Surface pressure is usually used as data to account for the atmospheric mass. For the GRACE mission, such assumptions bring non-negligible errors and are no longer valid [*Swenson and Wahr*, 2002]. In this chapter, the rigorous formulas to compute geopotential coefficients from the 3-D pressure field are introduced. Its difference with other approaches is quantified. In addition, we use alternative analysis outputs (ECMWF TOGA and NCEP operational analyses) to validate the current GRACE AOD1B product.

7.1 Equations relating pressure to gravity

The Earth's gravitational field is usually described by a set of spherical harmonic coefficients [*Heiskanen and Moritz*, 1967]. Since the atmosphere mass is enclosed within an envelope with the boundaries at places where the pressure is zero, we can write the potential ΔV caused by the atmosphere mass at an arbitrary position (r', θ', λ') outside the atmospheric envelope in the same form as the potential of the Earth's gravitational field, but normalized by the Earth's mean radius a and other constants (e.g., M_e).

$$\Delta V(r', \theta', \lambda') = \frac{GM_e}{r'} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \left(\frac{a}{r'}\right)^n P_{nm}(\cos\theta')(\Delta C_{nm} \cos m\lambda' + \Delta S_{nm} \sin m\lambda') \tag{7.1}$$

where

$$\begin{bmatrix} \Delta C_{nm} \\ \Delta S_{nm} \end{bmatrix} = \frac{1}{(2n+1)M_e a^n} \int_0^{2\pi} \int_0^{\pi} \int_{r_s}^{r_{top}} r^n P_{nm}(\cos\theta) \begin{bmatrix} \cos m\lambda \\ \sin m\lambda \end{bmatrix} dM \tag{7.2}$$

$$dM = \rho(r,\theta,\lambda)dV = \rho(r,\theta,\lambda)r^2 dr(\sin\theta d\theta d\lambda) = \rho(r,\theta,\lambda)r^2 dr dS \tag{7.3}$$

Substituting equation (7.3) into equation (7.2), results in

$$\begin{bmatrix} \Delta C_{nm} \\ \Delta S_{nm} \end{bmatrix} = \frac{1}{(2n+1)M_e a^n} \int_0^{2\pi} \int_0^{\pi} K(n) P_{nm}(\cos\theta) \begin{bmatrix} \cos m\lambda \\ \sin m\lambda \end{bmatrix} \sin\theta d\theta d\lambda \tag{7.4}$$

with

$$K(n) = \int_{r_s}^{r_{top}} r^{n+2} \rho(r,\theta,\lambda) dr \tag{7.5}$$

where in above equations:

G:	The gravitational constant
a:	The mean radius of the Earth (scale factor)
dS:	The surface element (unit area)
M_e:	The mass of the Earth
ρ:	The air density, a function of location
r_s:	The radius of the computation point at the Earth's surface
r_{top}:	The radius of the computation point at the top of the atmosphere
θ:	The geocentric co-latitude $\theta = 90^\circ - \phi$ of the computation point
λ:	The geocentric longitude of the computation point

The radial distance r (from the computational point to the geocenter, OP) is illustrated in Figure 7.1, which shows a cross section of the ellipsoidal Earth. The coordinates of the computational point in the plane are represented as [*Jekeli*, 2000]:

$$x = (N+h)\cos\phi \tag{7.6}$$

$$y = [N(1-e^2)+h]\sin\phi \tag{7.7}$$

where N is the radius of curvature of the point in the plane, h is the ellipsoidal height, ϕ is the geodetic latitude, and e is the eccentricity. Thus, the radial distance is:

$$r = \sqrt{(N+h)^2\cos^2\phi + [N(1-e^2)+h]^2\sin^2\phi} \tag{7.8}$$

An approximate expression for the radial distance r is often used in the literature [*Swenson and Wahr*, 2002; *Boy and Chao*, 2005]:

$$r = r_0 + \zeta(\phi,\lambda) + H(\phi,\lambda) = r_0 + h(\phi,\lambda) \tag{7.9}$$

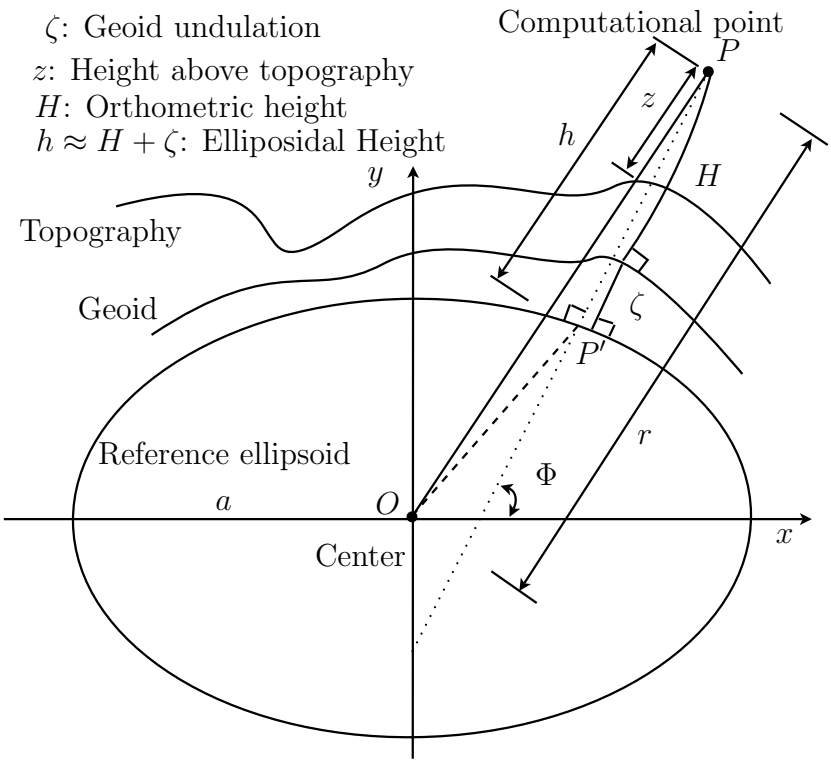


Figure 7.1: The distance from the computational point to the geocenter.

where ζ is the geoid undulation, and H is the orthometric height of the computational point. r_0 is the distance of OP', which is approximated by the Earth's mean radius a in the above references.

Usually, the air density is not explicitly given in the meteorological products. To evaluate equation (7.4), we need to derive the air density $\rho(r, \theta, \lambda)$ from the grid values of pressure, temperature, and humidity given in model analyses. We have two options to obtain this relationship. One is the equation of state of an ideal gas (ideal gas law), the other is the hydrostatic equilibrium.

7.1.1 Using the equation of state

The equation of state gives a relationship among pressure, temperature, and density.

$$P = \rho R T_v \qquad \text{or} \qquad \rho = \frac{P}{R T_v} \tag{7.10}$$

P is the pressure, R is the gas constant, T_v is the virtual temperature, and ρ is the air density. Substituting (7.10) into (7.4), yields,

$$\begin{bmatrix} \Delta C_{nm} \\ \Delta S_{nm} \end{bmatrix} = \frac{a^2}{(2n+1)M_e} \int_0^{2\pi} \int_0^{\pi} \left[\int_{r_s}^{r_{top}} \left(\frac{r}{a}\right)^{n+2} \frac{P}{RT_v} dr \right] P_{nm}(\cos\theta) \begin{bmatrix} \cos m\lambda \\ \sin m\lambda \end{bmatrix} \sin\theta d\theta d\lambda \tag{7.11}$$

where

$$r_s = r_0 + \zeta(\phi, \lambda) + H_s(\phi, \lambda) \tag{7.12}$$

H_s is the height of the surface topography. Normally, variables in meteorological products are defined on standard pressure levels, whose vertical position is given by their geopotential heights. The definition of the geopotential height and its relationship with the geometric height are described in detail in Appendix A, where we also compared various approximate methods. We choose the Smithsonian method because of its accuracy and computational simplicity. From equations (A.9) and (7.9), we obtain

$$r = r_0 + \zeta + \frac{R_f(\phi)\Phi}{\frac{g(\phi)}{g_0} R_f(\phi) - \Phi} \tag{7.13}$$

After replacing r in (7.11) by equation (7.13), we can calculate the spherical harmonic coefficients using 3-D gridded pressure, temperature, and geopotential height.

7.1.2 Using the hydrostatic equation

For synoptic scale atmospheric motions, the vertical accelerations in the dynamic equations are normally negligible. The atmosphere is in nearly hydrostatic balance [*Holton*, 1992]. The differential form of the hydrostatic equation for an air column is:

$$\rho(r, \theta, \lambda) dr = -\frac{dP}{g(\theta, r)} \tag{7.14}$$

or in 2-D form:

$$\sigma(\theta, \lambda) = -\frac{P_s}{g(\theta)} \tag{7.15}$$

P_s is the surface pressure and $\sigma(\theta, \lambda)$ is the surface density. For computational simplicity, the gravity g is only a function of co-latitude θ and height r in 3-D form and a function of co-latitude θ in 2-D case. Substituting (7.14) into equation (7.5), gives us:

$$K(n) = -\int_{P_s}^{0} r^{n+2} \frac{dP}{g(\theta, r)} \tag{7.16}$$

The r in the above equation can be replaced by equation (7.13). Finally, considering the elastic deformation of the solid Earth, equation (7.4) becomes:

$$\begin{bmatrix} \Delta C_{nm} \\ \Delta S_{nm} \end{bmatrix} = -\frac{a^2(1+k_n)}{(2n+1)M_e} \int_0^{2\pi}\int_0^{\pi} \left[\int_{P_s}^{0} \left(\frac{r}{a}\right)^{n+2} \frac{dP}{g(\theta,r)} \right] P_{nm}(\cos\theta) \begin{bmatrix} \cos m\lambda \\ \sin m\lambda \end{bmatrix} \sin\theta d\theta d\lambda \tag{7.17}$$

where k_n is the load Love number for degree n [*Farrel*, 1972].

7.1.2.1 Layered pressure on real Earth's surface

In reality, the surface pressure is not a complete measure of the atmospheric mass [*Trenberth*, 1981]. Due to the curvature of the Earth, the vertical atmospheric columns become wedges with inclined lateral surfaces. The mass of every wedge is balanced by not only the pressure at the bottom of the surface but also the support to its lateral surfaces [*Rannon et al.*, 1997]. As a result, the product of the surface pressure and the bottom area is always less than the overlaid air weight above that area. In the 3-D case, the equation (7.16) can be re-written as:

$$K(n) = -\int_{P_s}^{0} (r_s + z)^{n+2} \frac{dP}{g(\theta, r_s + z)} \tag{7.18}$$

with

$$r_s = r_0 + \zeta + \frac{R_f(\theta)\Phi_s}{\frac{g(\theta)}{g_0}R_f(\theta) - \Phi_s} \tag{7.19}$$

z is the height above the Earth's topographic surface. Replacing $K(n)$ in equation (7.4) by (7.18), we have a formula that converts the layered (3-D) pressure field to geopotential coefficients.

The meteorological products are given either on isobaric levels (equal pressure) or hybrid half levels. For the product provided on isobaric levels, the goepotential heights are explicitly provided, the layered temperature and humidity data are used for the computation of the virtual temperature. For the products provided in hybrid coordinates, the temperature and humidity data are given at half levels. The coordinates define the edge of the atmospheric layers in terms of the surface pressure.

$$P_{i+\frac{1}{2}} = a_{i+\frac{1}{2}} + b_{i+\frac{1}{2}} P_s \tag{7.20}$$

$a_{i+\frac{1}{2}}$ and $b_{i+\frac{1}{2}}$ are model dependent coefficients for every half level. The subscript $(i+\frac{1}{2})$ indicates that the variable is at the half level. Using the digitized hypsometric equation [*Holton*, 1992], the geopotential height at each half level is given by:

$$\Phi_{i+\frac{1}{2}} = \Phi_s + \frac{1}{g_0} \sum_{n=i+1}^{n=N_{top}} R_{dry} T_v \ln \frac{P_{i+\frac{1}{2}}}{P_{i-\frac{1}{2}}} \tag{7.21}$$

where Φ_s is the surface orography, R_{dry} is the gas constant for dry air and T_v is the virtual temperature and can be computed by equation (2.32)

Under different assumptions of the atmospheric structure and approximations of reference surfaces, equation (7.18) can be simplified to different forms. The top standard pressure level is approximately 50 km above the MSL, which is very small compared to the radius of the Earth. Using equation (A.2) for g, the $r_s + z$ term in (7.18) can be approximated by its first order expansion:

$$K(n) = -\frac{(r_s)^{n+2}}{g(\theta)} \int_{P_s}^{0} \left(1 + (n+4)\frac{z}{r_s}\right) dP \tag{7.22}$$

The above equation can be re-organized as:

$$K(n) = \frac{(r_s)^{n+2}}{g(\theta)} \left[P_s - \int_{P_s}^{0} (n+4)\frac{z}{r_s} dP\right] \tag{7.23}$$

The second term in the bracket (a negative value) represents the missing part if only the surface pressure is used in the computation. To quantify the effect of this term, we use the U. S. Standard Atmosphere [*U. S. Standard Atmosphere*, 1976] to calculate the integral. By letting $n = 0$, and $r_s = 6371$ km, the magnitude of this term is about 4.7 hPa, and it increases as n becomes larger.

7.1.2.2 Surface pressure on a spherical Earth

We can neglect the thickness of the atmosphere ($z = 0$) and assume that the pressure only concentrates on a thin layer above the Earth's surface approximated by a reference sphere with a radius a (neglecting the oblateness of the Earth and real surface topography). Let $r_s = a$ and $g(\theta) = g_0$, equation (7.23) is simplified to:

$$K(n) = a^{n+2} \frac{P_s(\theta, \lambda)}{g_0} \tag{7.24}$$

The final formula is:

$$\begin{bmatrix} \Delta C_{nm} \\ \Delta S_{nm} \end{bmatrix} = \frac{a^2(1+k_n)}{(2n+1)M_e} \int_0^{2\pi} \int_0^{\pi} \left[\frac{P_s(\theta, \lambda)}{g_0}\right] P_{nm}(\cos\theta) \begin{bmatrix} \cos m\lambda \\ \sin m\lambda \end{bmatrix} \sin\theta d\theta d\lambda \tag{7.25}$$

This expression can be computed using the FFT method.

7.1.2.3 Surface pressure on the real Earth's surface

The surface topography is usually described as the orthometric height above the Mean Sea Level. If we only neglect the thickness of the atmosphere, equation (7.23) becomes:

$$K(n) = r_s^{n+2}(\theta, \lambda) \frac{P_s(\theta, \lambda)}{g(\theta)} \tag{7.26}$$

Thus, equation (7.17) is simplified to:

$$\begin{bmatrix} \Delta C_{nm} \\ \Delta S_{nm} \end{bmatrix} = \frac{a^2(1+k_n)}{(2n+1)M_e} \int_0^{2\pi} \int_0^{\pi} \left(\frac{r_s}{a}\right)^{n+2} \frac{P_s(\theta,\lambda)}{g(\theta)} P_{nm}(\cos\theta) \begin{bmatrix} \cos m\lambda \\ \sin m\lambda \end{bmatrix} \sin\theta d\theta d\lambda \tag{7.27}$$

Equation (7.18) becomes similar to formula (3-18) given in the GRACE AOD1B description document [*Flechtner*, 2005] if we assume $g(\theta) = g_0$ and $R_f(\theta) = a$. In the document, they also use the surface geopotential height Φ_s to approximate the geoid undulation ξ. The difference of these two values sometimes is very big. In addition, their method ignores the dependence of the gravity on latitudes.

7.1.3 Calculation procedures and the mean field

The datasets used include the ECMWF TOGA operational analysis and NCEP final (FNL) operational analysis. They are both available 4 times daily. ECMWF TOGA is in the spherical harmonic format and has a horizontal resolution of T106. Vertically, it has 21 standard pressure levels from 1000 hPa to 1 hPa. Before the computation, we first extract the 3-D geopotential height, temperature, specific humidity and surface orography from the dataset, and then convert them to a 3-D regular $1^\circ \times 1^\circ$ grid. The NCEP dataset is converted to a $1^\circ \times 1^\circ$ regular grid from a much higher resolution. Vertically, it has 26 mandatory pressure levels from 1000 hPa to 10 hPa. We extract the same 3-D variables as in ECMWF from the original file to save storage space.

To evaluate the atmospheric variations, a long-term (usually longer than one year) mean is subtracted from pressure fields. For the formulas using 2-D surface pressure, it is relatively easy to compute the mean of the surface pressure. For the 3-D case, the computation is intensive. The integral (7.18) has to be calculated for each grid and each degree. If we want to compute a 3-D yearly mean, there are 365×4 3-D fields. For each field, there are 360×180 vertical columns. For each air column, we need to compute the integral (7.18) for each degree (< 100).

A mean field for the year 2001 was prepared at the beginning. For surface pressure, we do not directly use the surface pressure provided in the datasets. Instead, we calculate the surface pressure for both ECMWF and NCEP using the same topography based on their own 3-D fields, which guarantee that the two surface pressure fields are referenced to the same topography. For each grid, 4 corner points are used to compute a block mean. The geopotential coefficients are calculated by the numerical integration method (See Appendix D) after the subtraction of the mean field.

7.2 Computational results

In order to remove the short term variations in the atmosphere and ocean, GFZ has been preparing the Atmosphere and Ocean De-aliasing (AOD) product, which is released as AOD1B. This product is based on the ECMWF Integrated Forecast System (IFS) at synoptic time 00Z, 06Z, 12Z, 18Z (Z stands for Greenwich time). The horizontal resolution of the data is T213/N160, which corresponds to about 0.56°. Vertically, the data are given on 60 layers from the surface to 0.1 hPa. We summarize the datasets used in this study in Table 7.1.

Dataset	Interval	Horizontal	Vertical	Level type	Representation
ECMWF (AOD1B)	6 hours	T213/N160 approx. 0.56°	60 layers from surface to 0.1 hPa	hybrid levels	spherical harmonics
ECMWF TOGA	6 hours	T106/N80 1.125°	21 layers from 1000 hPa to 1 hPa	pressure levels	spherical harmonics
NCEP FNL	6 hours	1°× 1°	26 layers from 1000 hPa to 10 hPa	pressure levels	regular grid

Table 7.1: The meteorological data used in the study

It is apparent that GFZ uses much higher resolution data than we do. Our dataset (pressure level) is derived from the original dataset (hybrid level) for diagnostic purposes, but with a lower resolution. It is obtained from the National Center for Atmospheric Research (NCAR) data center. The original ECMWF data is not available to us. In order to validate our computation, we first compare our results with the GFZ AOD1B product.

7.2.1 Comparison with the GRACE AOD1B products

Two methods are introduced to obtain the air density in the formula which converts pressures to geopotential coefficients. One is the equation of state formula for ideal gases, the other is the hydrostatic equation. We use the hydrostatic equation to compare with the AOD1B products, since it uses the hydrostatic formula as well. However, we adopt more precise formulas in the computation, e.g.,

- A better formula in the conversion of geopotential height and geometric height;

- The latitudinal variation of the gravity is considered;
- The exact ellipsoidal Earth radius is computed;
- Real geoid height is used in the formula (AOD1B approximates the geoid height using the surface topography).

Some of these factors may or may not affect the accuracy of the final results, this is assessed in this study.

Figure 7.2 shows the degree amplitude for a whole month of AOD1B product and our hydrostatic computation in April, 2003. The expected GRACE measurement error is plotted in the figure to indicate the GRACE sensitivity. The monthly mean differences are below the GRACE sensitivity except the first several degrees. Considering the fact that the actual GRACE sensitivity is worse than this ideal case, we might conclude that their differences are below the GRACE sensitivity. The RMS difference (defined as the RMS of their differences) is also below the GRACE sensitivity except for degree 2, which might be caused by the latitudinal variation of the gravity or the different model resolutions. These differences are also plotted in a geographic map in Figure 7.3. The mean difference is relatively larger in the polar region and a clear zonal feature is depicted in the RMS map.

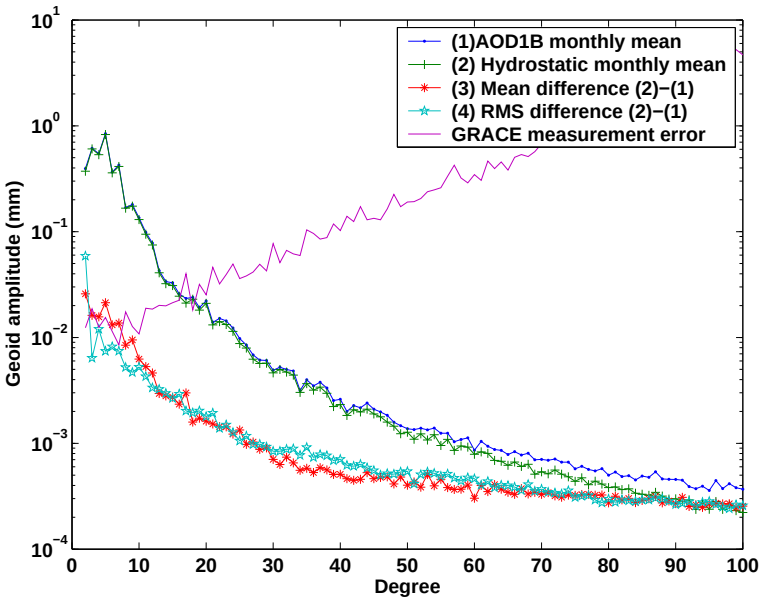
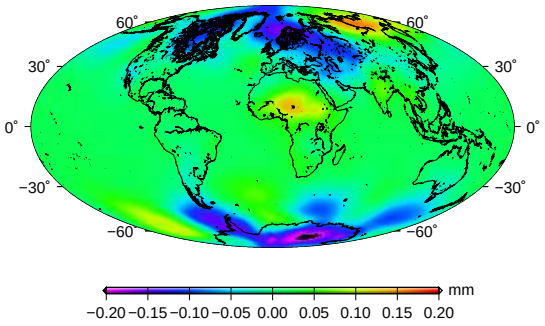
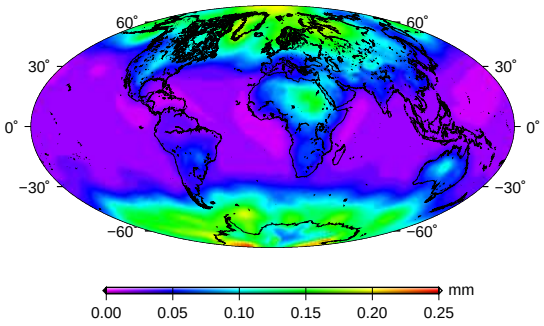


Figure 7.2: Spectrum of the AOD1B product and our hydrostatic computation for April, 2003.



(a) Mean difference



(b) RMS difference

Figure 7.3: Mean and RMS differences between the AOD1B product and our computation.

It is surprising to see that the difference of these two computations are below the GRACE sensitivity. The approximation of the geoid by the topography adopted in the GFZ AOD1B product has little effect on the results. We suspect that a large part of the differences in the algorithms are canceled by their respective yearly mean fields. Since we do not have the original ECMWF data to produce the AOD1B product, in the following comparisons, it is replaced by our ECMWF data but with a similar algorithm to AOD1B (We also call this AOD1B).

7.2.2 Difference between 3-D computations

7.2.2.1 Hydrostatic vs. AOD1B

To verify the effect of the mean field, we compute the spherical harmonic coefficents from the same ECMWF data using our formula and the AOD1B formula. For each method, two sets of 30-day coefficients are calculated. A yearly mean is only removed from one set of coefficients. The coefficients with and without their respective means removed are compared in Figure 7.4. It is shown that the monthly mean difference with mean removed is almost below the sensitivity, while the difference without removing their respective means is more than one order of magnitude larger than the other, but their RMS difference is very small except for degree 2. This implies that the two mean fields for our hydrostatic and AOD1B computations are different. The effects of the topography, latitude dependent gravity and other aspects commonly exist in both monthly and yearly mean fields. These effects are mostly canceled after subtracting the mean field. This would explain why the results are similar, despite of the differences in their assumptions.

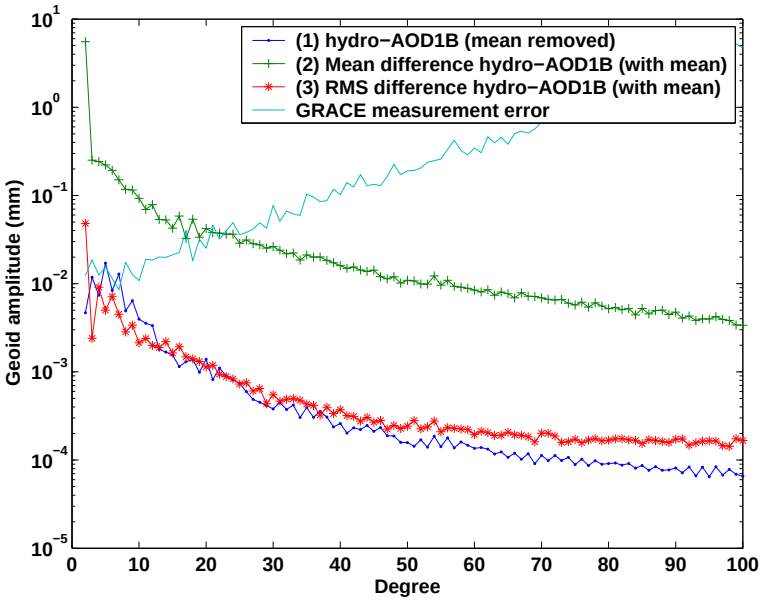
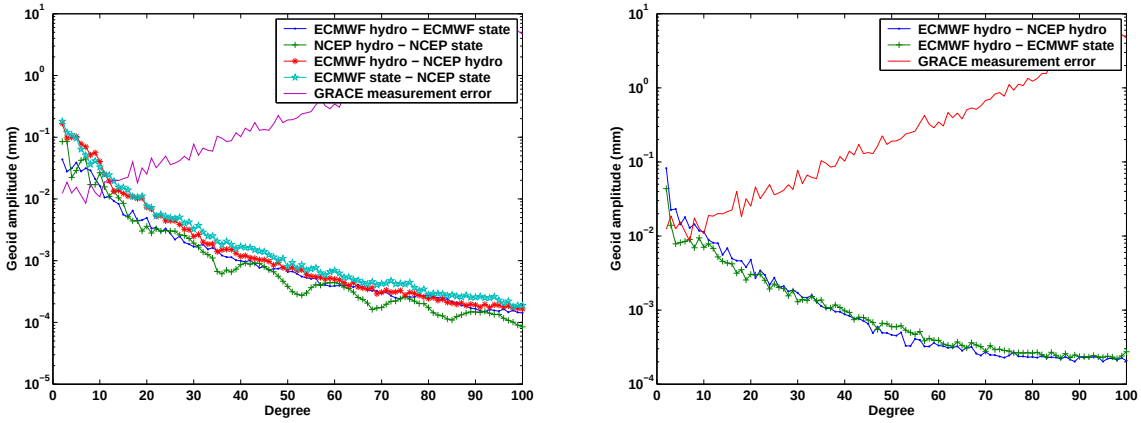


Figure 7.4: Spectra for the comparisons of hydrostatic and AOD1B computation with and without mean removed.

7.2.2.2 3-D Hydrostatic equation vs. Equation of state

The air density can be derived from the hydrostatic equation or the equation of state. One can use either equation (7.11) or (7.17) to calculate the geopotential coefficients. In Figure 7.5, the spectra for the mean and the RMS difference between the coefficients computed using the equation of state and the hydrostatic equation are plotted. All the curves in the figure are above the GRACE sensitivity at around degree 12. It is shown that using the same dataset with different methods as well as using the same method with different models brings the same level of difference. There is no clear explanation for the differences between the two methods. Theoretically, they should bring the same result. In reality, however, 3-D datasets may not strictly obey these two equations, i.e., the air is not in perfect hydrostatic balance and/or does not strictly follow the ideal gas law. In the hydrostatic equation (7.17), no temperature is required for the integration, while in the equation of state formula (7.11), no gravity information is needed for the integration. These effects cannot be removed by subtracting the mean field.

The spatial differences are shown in Figure 7.6. The monthly mean difference between the hydrostatic and the equation of state is larger over Antarctica for both ECMWF and NCEP. The magnitude of ECMWF is relatively smaller. These two equations are not consistent for the fields over Antarctica. (c) and (d) illustrate again



(a) Mean difference

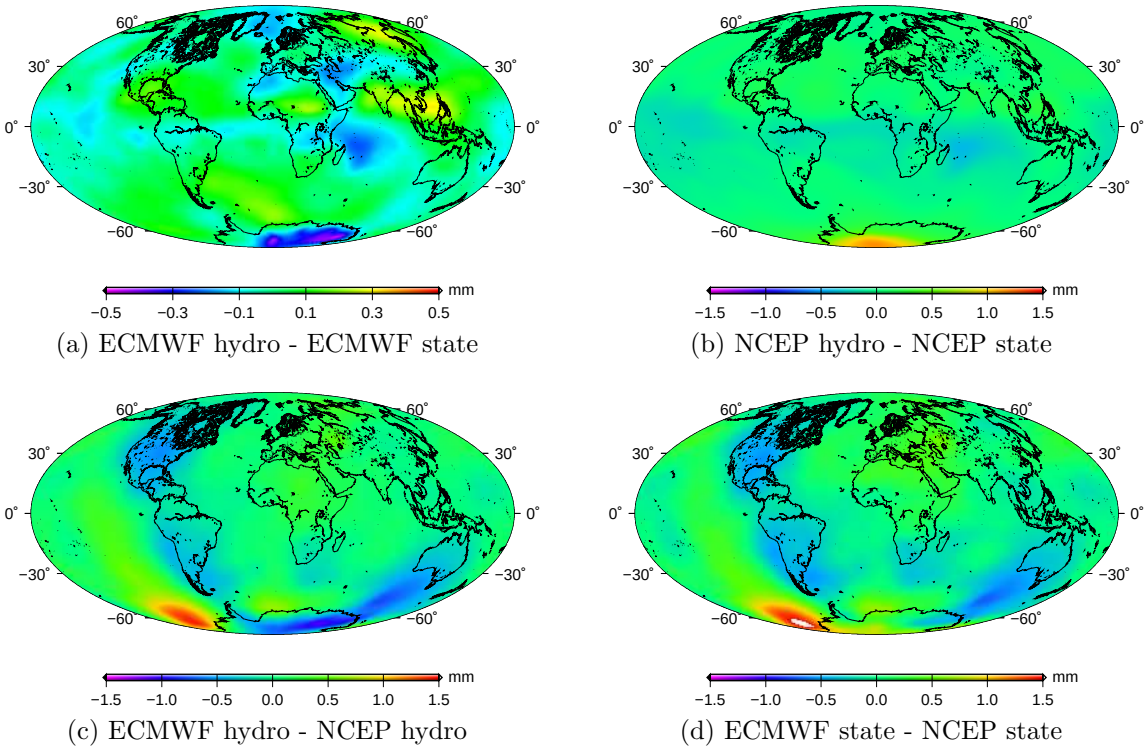
(b) RMS difference

Figure 7.5: Spectra for the montly mean and the RMS difference between the coefficients computed using the equation of state and the hydrostatic equations in April, 2003.

the large difference between ECMWF and NCEP at Southern Ocean region. It is still a question of which equation gives a better result. This is under investigation.

7.2.3 3-D computation vs. 2-D computation

Sometimes, it is convenient to use the surface pressure to compute the mass variations of the atmosphere by neglecting its vertical thickness. Equation (7.25) assumes that Earth is a sphere and equation (7.27) considers the roughness of the Earth's topography. By using the surface pressure, computation loads can be greatly reduced, and the FFT method can be applied to further improve the computational efficiency. Figure 7.7 shows the monthly mean difference between the 3-D computation and the 2-D spherical Earth assumption. The differences between 3-D hydrostatic results and 2-D spherical results are smaller than the expected GRACE sensitivity for both ECMWF and NCEP, which was not expected based on studies in some previous publications [e.g. *Swenson and Wahr*, 2002]. The differences using the equation of state are above the sensitivity around degree 12. For the comparisons of the 3-D computation with the 2-D realistic Earth assumption in Figure 7.8, the results are similar. The 2-D results do not suffer from neglecting the vertical atmospheric structure. The spatial distribution of these difference is plotted in Figure 7.9. For the figures related to the 3-D hydrostatic comparisons (a, b, e), there is almost no difference in the tropical region. The discrepancies concentrate in the middle and high latitudes. The southern ocean area shows relatively larger differences. Similar features are not



(a) ECMWF hydro - ECMWF state

(b) NCEP hydro - NCEP state

(c) ECMWF hydro - NCEP hydro

(d) ECMWF state - NCEP state

Figure 7.6: The spatial differences for the coefficents computed using two models and two methods.

present in the 3-D state figures (c, d, f). The magnitude of the difference is larger than that of the 3-D hydrostatic comparisons. The difference exists in the tropical region as well, but is dominant in the Antarctica region.

In order to investigate whether those effects are canceled by their respective yearly mean fields, we re-compute the coefficients without removing their yearly means. Figure 7.10 shows that there are biases between the 3-D and 2-D computations when leaving the mean field intact, and the RMS of their differences are below the sensitivity except for degree 2. The 2-D realistic assumption does bring a smaller bias when compared to the 3-D computation.

7.2.4 Summary

New gravity missions require an accurate modeling of the atmospheric mass change. We introduce rigorous formulas to convert the 3-D meteorological fields (pressure, temperature, and humidity) to geopotential coefficents. We use a more precise formula to convert the geopotential height to geometric height. The oblateness of the

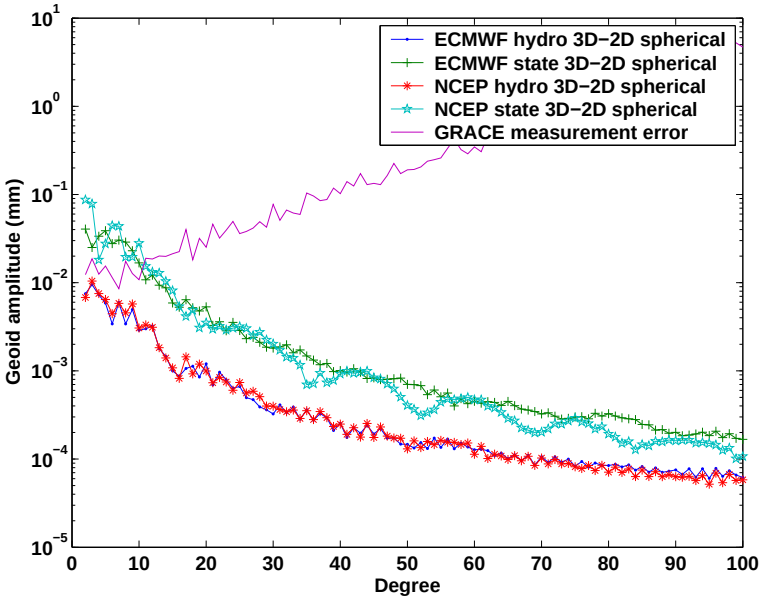


Figure 7.7: Monthly mean difference between 3-D computations and 2-D spherical Earth in April, 2003.

Earth and the latitude dependence of gravity are also carefully taken care of. The required air density can be derived either by the hydrostatic equation or the equation of state. With some approximation, the 3-D formula can be simplified to a 2-D version, which only requires the surface pressure. Using surface pressure only can greatly reduce the computational load.

It is surprising to find that the differences between the 3-D computation using the hydrostatic equation and the 2-D formulas are below the expected GRACE sensitivity after removing their respective mean fields, which disagrees with a previous study [*Swenson and Wahr*, 2002]. We further prove that the effects of the vertical atmospheric structure, topography, and the oblateness of the Earth can be largely compensated by subtracting the mean field for both ECMWF and NCEP data. Consequently, we could argue that the surface pressure is adequate for the atmospheric de-aliasing purposes. However, the comparison between the 3-D hydrostatic formula and the 3-D equation of state formula shows slightly different results. The exact reason for this discrepancy is not known yet. We suspect that the consistency between the data and the model might cause the problem. In a previous study, *Boy and Chao* [2005] showed agreement between these two formulas on a high resolution hybrid level dataset, but the authors reported disagreement between the hybrid level data and pressure level data based on the same hydrostatic equation. This contradicts our analysis. The difference between our results and AOD1B products using

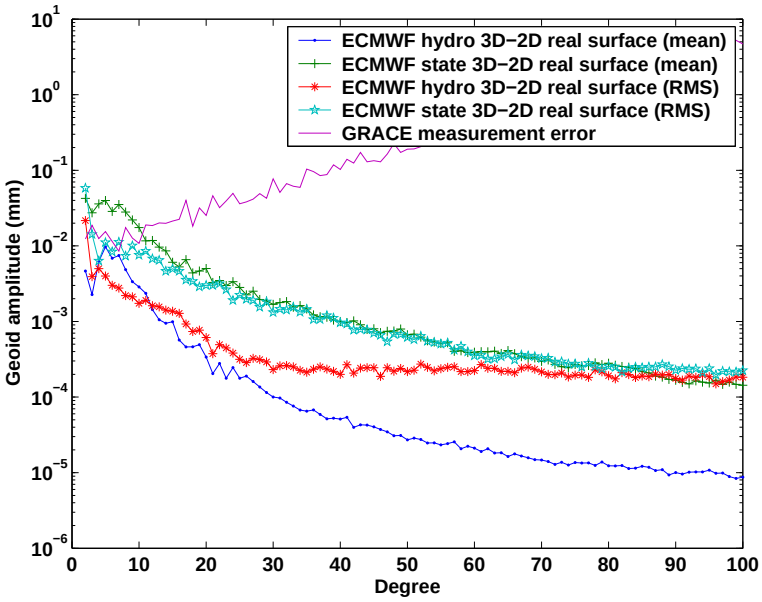


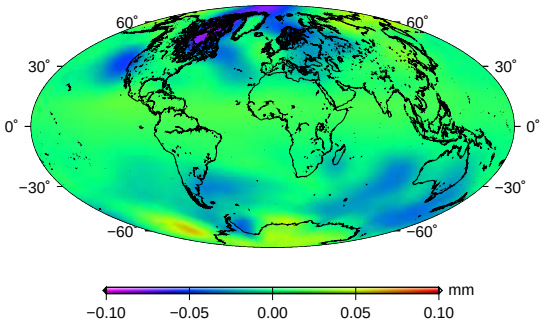
Figure 7.8: Mean and RMS difference between 3-D computations and 2-D realistic Earth.

the hydrostatic equation is almost below the theoretical GRACE sensitivity except for the first several degrees. We have addressed the fact that there are differences in the detailed implementations, and our dataset consists of low resolution pressure level data. The AOD1B product was produced by a high resolution hybrid level data. We suspect the dataset used may also affect the conclusion of the analysis.

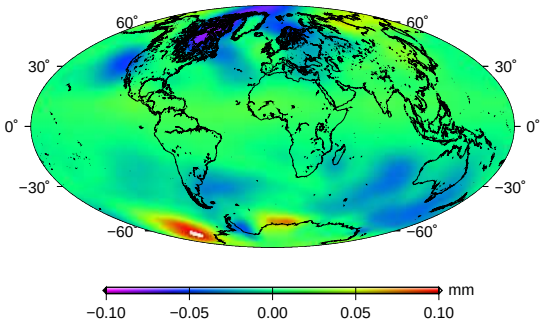
The difference between ECMWF and NCEP, once again, reveals the discrepancy between these two models, especially over the Antarctic region. Their difference cannot be canceled by removing their respective mean fields, which indicates that the atmospheric variations are modeled differently by these two models.

Usually, it is hard to know the exact error in the 3-D fields. Surface pressure can be validated by ground station measurements, like we did with AWS. For 3-D fields, comparisons with radiosondes or radio occultation profiles can reveal some of the story. In those measurements, however, there also exists errors, which makes the comparison a difficult task. Fortunately, the NWP models are keep improving in precision, and horizontal and vertical resolutions. Especially the improvements over southern latitudes are noticeable by assimilating more satellite measurements.

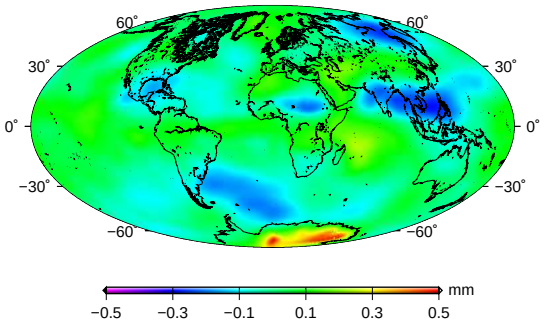
For the areas with large hydrological signals, omitting of the vertical structure of the atmosphere brings negligible error. The surface pressure is adequate for those applications.



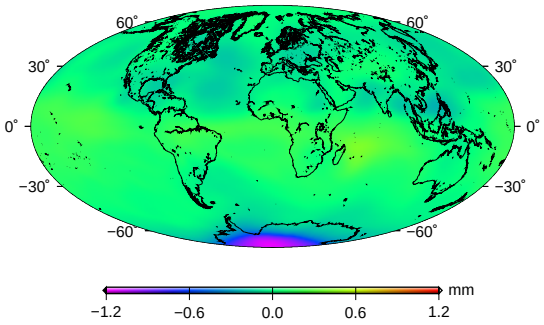
(a) ECMWF 3-D hydro - 2-D spherical



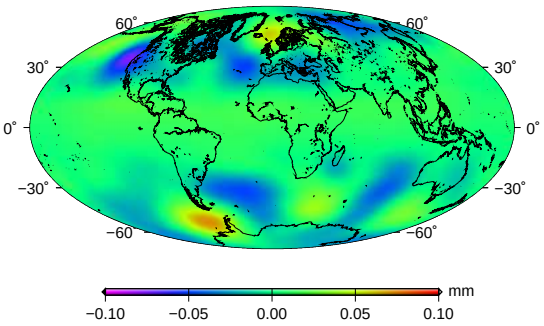
(b) NCEP 3-D hydro - 2-D spherical



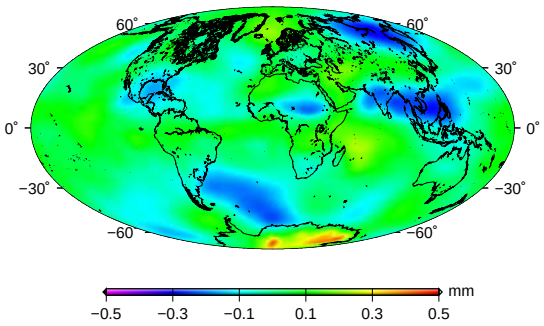
(c) ECMWF 3-D state - 2-D spherical



(d) NCEP 3-D state - 2-D spherical



(e) ECMWF 3-D hydro - 2-D real surface



(f) ECMWF 3-D state - 2-D real surface

Figure 7.9: The spatial difference for the coefficients computed using 2-D and 3-D methods.

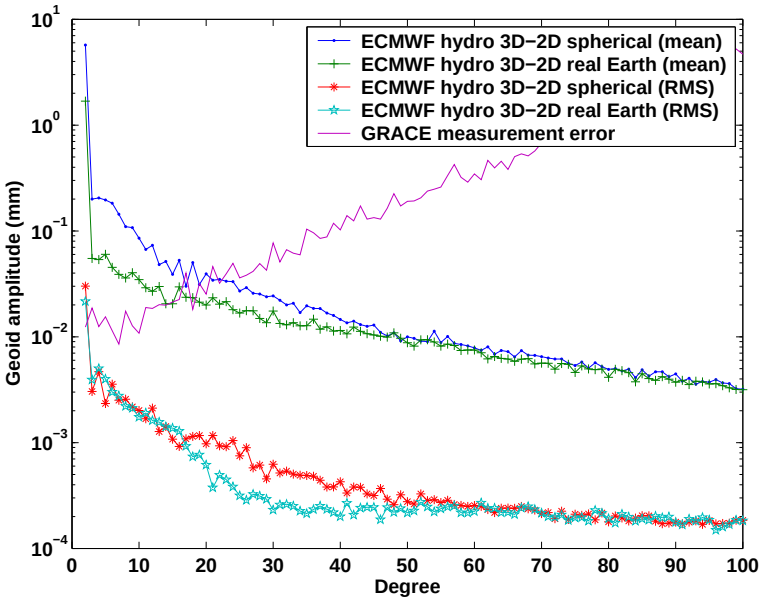


Figure 7.10: Monthly Mean and RMS difference between 3-D and 2-D computations with the yearly mean un-removed.

CHAPTER 8

CONCLUSIONS AND OUTLOOK

8.1 Summary and conclusions

Atmospheric mass is only about one millionth of the mass of the solid Earth, but it has the largest temporal variation compared to the hydrological, cryospheric, oceanic and solid earth signals. The amplitude of its variation is similar to that of the hydrological signal. Unlike the tidal signals, the atmospheric variation has wide temporal and spatial spectra. Modeling such a signal and determining its error distribution become very complicated. Satellite gravity missions such as GRACE provide an unprecedented opportunity to sense those hard-to-measure signals in remote regions as mass constraints for geophysical models, but they require an accurate pressure model to separate other interesting signals. Antarctica represents one such regions with many mysteries undiscovered due to its remote location and extreme environment. Since NWP models largely depend on the accuracy and the distribution of *in-situ* observations, atmospheric pressure from NWP model has the biggest uncertainty in the Southern Ocean and Antarctic regions. The atmospheric errors increase the uncertainty of the recovered integrated mass signal, and prevent a clear separation of the contributions from each of the Earth's sub-systems. Quantifying the error budget from the atmospheric contribution and improving the current NWP modeling in Antarctica become very important tasks.

In this research, we investigated the potential of using GPS occultation pressure profile to improve the Antarctic pressure modeling. We paid special attention to the characteristics of GPS occultation measurement in the polar regions. A software system has been developed to retrieve the pressure, temperature, and humidity profile from excessive delay measurements based on the geometrical optics and radio holographic methods. By analyzing results from our software, we confirmed that the occultation measurement has the best resolution in the vertical range around 5-30 km for pressure and other variables. In the lower troposphere especially in the tropics, the errors become relatively large because the measurements suffer from the rapidly varying moisture content and receiver tracking problem. Various methods to deal

with this problem have been introduced in Chapter 2 in the "software" and "algorithm" aspects. A complete solution requires the change of the GPS "hardware" tracking to an "open loop algorithm" [*Sokolovskiy*, 2004]. In the polar regions, the multipath problem is alleviated and becomes less critical because that the extremely cold temperature makes the air relatively dry and hardly holds moisture. Our results show that occultation measurements have higher SNR and reduced noise than in the tropics. Consequently, most of the profiles in polar regions can penetrate down to less than 1 km above the ground surface. The profile retrieval algorithms are based on various assumptions, simplifications, smoothing and initializations. To test the sensitivity of the final retrieval results to factors like algorithm, time, satellite, and area of interest, we examined the CHAMP and SAC-C profiles from UCAR and JPL with the interpolated ECMWF and NCEP pressure profiles at 10 km altitude (less affected by moisture), and chose a period that the two LEOs used a similar version of tracking firmware and all the profiles have passed the quality control. We found that they generally agree with each other in most of the regions, but there are discrepancies possibly caused by different retrieval algorithms and satellite signal tracking in some regions. Southern Ocean and Antarctica have the largest disagreement, where the difference can be as large as 5 hPa. Individual profile needs to be examined to pinpoint the cause of the difference. NWP models usually have distinct regional performance primarily due to the unevenly distributed observations. We chose to validate the CHAMP profiles against ECMWF, NCEP and radiosonde data in 5 geographic regions. It is found that the regional averaged pressure bias and the standard deviation difference are still large at or near the Earth surface, especially over the southern polar region (> 3 hPa). CHAMP derived pressures is systematically smaller compared to other data in all altitudes (1-30 km) in the southern polar region. The regional averaged ECMWF and NCEP do agree well in the profile locations, with a slightly larger standard deviation difference only in the southern polar region. This depicts the internal consistency between NWP models; on the other hand, it also implies that common errors may exist in both models. Identifying whether the bias exists in the model or comes from the occultation profile is very meaningful. It can be used to improve either the NWP model or occultation depending where the error resides, although it is a difficult task in such a data-sparse region.

To avoid the effects of various assumptions, errors and approximations in the final occultation retrieval results, we have the option to assimilate the quantities in the early stage of the occultation retrieval procedure directly to NWP models. These variables can be the high-level pressure, temperature and moisture profiles, middle-level refractivity profiles, and low-level bending angles or excessive phase measurements. Assimilating high-level data is relatively easy, but it suffers from all the errors in the algorithm. While for the low-level data, the non-linear relationship makes the formulation for the forward and adjoint modeling very complicated. In this study, we

adopted a 1-D variational approach and chose CHAMP refractivity profiles as observation (relatively cost effective) to combine with a background model from ECMWF. Our 1DVar results show remarkable improvement of the temperature in 200-300 hPa region. In addition, the high vertical accuracy of the occultation measurements in the upper atmosphere can be propagated down to have positive impact (error reduced to 70% of the original) on the surface pressure though the hydrostatic relation. This demonstrates the potential of the occultation measurements for the improvement of pressure field. After eliminating the existing limitation of the occultation, the increased number of observations and the reduced multipath effect make the occultation a very promising technique to improve the weather modeling in the Antarctic region.

The pressure errors and uncertainties embedded in NWP models degrade the GRACE recovered gravity change. The anticipated pressure error should be under 1 hPa. We examined the model uncertainties and found that the discrepancy can be as large as 6-8 hPa around the Southern Ocean and the Antarctic regions on daily and monthly time scale. Topographical effects need to be taken care of when NWP models adopt different terrain models. The effects of topography can be reduced by removing a long-term mean field. After converting pressure fields to spherical harmonic coefficients, the degree 0 and 1 terms are neglected by assuming that the total mass is conserved and the coordinate center is at the center of mass. Barometric (thermal) tides are not negligible and are usually well treated in the ocean tide model. The difference between ECMWF and NCEP is currently lower than the calibrated error of the GRACE monthly solution, but it contributes to the total error budget. We separate the total atmospheric errors into (1) the mis-modeled or un-modeled error that can be correctly sensed by GRACE; (2) the aliasing error caused by the imperfect modeling of the variability. The exact magnitude of (1) is hard to assess unless independent *in-situ* ground truth is available. The contribution from (2) can be simulated using an exact satellite orbit configuration with one model as truth and the other as nominal de-aliasing model. After de-aliasing using NCEP re-analysis data, we found that the middle to high degree spherical harmonic coefficients have noticeable (3-7 times smaller) improvements. Most of the long wavelength errors are well captured by GRACE. Their variations induce extra errors for the middle to high degree coefficients. Simulation using recent ECMWF and NCEP operational fields gave reduced aliasing error, indicating that the variations of the two models becomes similar and NCEP has improved skills. The IB assumption further decreases the amplitude of the aliasing error. The RMS pressure difference over land is also slightly changed for the IB case. We argue that the IB assumption makes the aliasing estimation too optimistic. In addition, the simulated aliasing error also depends on the models chosen in the simulation and may not represent a realistic distribution. In addition, we also show that aliasing is not a local phenomenon. The un-modeled pressure variations in Antarctica produce aliasing errors over the globe. Improving Antarctic pressure modeling can help to reduce the global aliasing error.

The NWP models was validated against the AWS data in Antarctica. Using the 10-min sampling interval surface pressure measurements, we demonstrated that the 6-hour analyzed model data can capture most of the power by a spectral analysis. Linear interpolation is applicable for obtaining high sample-rate pressure data. In the validation, ECMWF exhibits a much better agreement with the AWS observations in terms of standard deviation than NCEP re-analysis. The not-well-surveyed AWS station elevation increases the pressure bias, although the bias is not always a constant value throughout the whole year. Part of the biases can be removed by subtracting a long-term mean. The standard deviation differences reveal that larger disagreements tend to happen during winter months, in the coastal regions where increasing weather activities occur, and in regions with sharp elevation change. The standard deviation in the high-elevation interior is generally lower than 2.0 hPa, and is larger than 2.0 hPa around the coastal region. NCEP re-analysis depicts a much larger disagreement. The comparison result between ECMWF and AWS might bring an optimistic result, since some AWS observations might have already been assimilated into ECMWF. The errors revealed in this study can serve as a realistic low bound for the model errors, provided that there is little error in the AWS observations. However, the AWS observations are not always reliable due to the harsh weather conditions and poor maintenance. We also found that the error distribution between the AWS and NWP models is different from the difference between two NWP models. The AWS data show that the errors are large along the coastline, while NWP model difference is large in the interior of Antarctica, which implies that some of the simulations using the model difference may not reflect the realistic error distributions.

In the final chapter, we investigated the influences of different algorithms and assumptions of 2-D or 3-D atmospheric structures on the atmospheric de-aliasing product. Calculating the air density by the hydrostatic equation or the equation of state brings non-negligible difference. The exact reason is at present unknown. To generate our own AOD product, we proposed several improvements compared to the GFZ's AOD1B algorithm. Meanwhile, we use a relatively low resolution ECMWF analysis data. It is found that the difference between these two AOD products is almost below the expected GRACE sensitivity. Surprisingly, we also found that the difference between the 3-D hydrostatic method and 2-D method is below the GRACE sensitivity. It is discovered later that the atmospheric structure and the latitudinal variation of gravity are largely compensated by removing their respective long-term means. If this is true, the computational load and storage space for atmospheric de-aliasing model can be greatly reduced by using the simplified 2-D method (surface pressure only). Removing the mean fields does not help to reduce the difference between the ECMWF and NCEP. There are still some fundamental differences between these two models. If computational burden is not a concern, using our improved model should bring a better result. Validation of the 3-D fields (temperature, pressure and humidity) is not as easy as the comparison using those surface observations.

These 3-D observations are hardly available in Antarctica. GPS occultation can be a good candidate in this situation.

8.2 Discussions and outlook

We have shown that accurate atmospheric modeling is important for reducing the total errors in GRACE gravity estimation. Improving not only the atmospheric modeling but also the ocean modeling as background models for de-aliasing the high-frequency variations is expected to have some significant impacts on the GRACE gravity solution, especially over the Antarctic region. Simulations based on model difference may incorrectly reflect the real error distribution. Sometimes, these errors are underestimated because of the correlation of the data and models. In addition, due to the nature of the GRACE orbital sampling, the aliasing of the interesting signals also needs to be well treated.

The NWP operational models are continuously improving their skills over southern latitudes [*Pendlebury et al.*, 2003] with better physical and dynamic modeling and by incorporating new satellite observations. Monitoring the change of NWP models can reduce the artificial errors introduced by the new data, e.g. the trends introduced by adding new satellite observations [*Hines et al.*, 2000; *Bromwich and Fogt*, 2004]. The new released GRACE solutions show a lower calibrated error after updating the onboard tracking firmware and background models. Further reducing the error is anticipated to achieve its baseline performance. At present, only the largest basins with large signal amplitude can be well recovered.

The assimilation of radio occultation measurements and its impact to the NWP models are restricted by the limited number of occultation data from current available satellites. The GPS occultation retrieval algorithm and onboard tracking software are being continuously improved for the lower troposphere. With the launch of the 6-satellite COSMIC constellation, the polar and other regions of the world will be more densely covered. A major improvement of the atmospheric modeling in Antarctica is expected. A re-processing of GRACE data with a most recent NWP model will increase the accuracy and resolution of gravity solutions, with reduced errors caused by the unknown atmospheric variability.

APPENDIX A

CONVERSION BETWEEN THE GEOPOTENTIAL HEIGHT AND THE GEOMETRIC HEIGHT

In most meteorological products, the pressure serves as the "natural" vertical coordinate. Normally, geopotential heights are given at the standard pressure levels, or it can be derived from hybrid model levels. The geopotential height Φ is defined as:

$$\Phi(H,\phi) = \frac{1}{g_0}\int_0^H g(H',\phi)dH' \tag{A.1}$$

where g_0 is defined as a constant value ($9.80665\ ms^{-2}$), which was originally the gravity at the $45\,^\circ$ latitude set by World Meteorological Organization (WMO). However, as gravity models continually improve, this value is slightly different from the current gravity at $45\,^\circ$. $g(H',\phi)$ is the local gravity (latitude and height dependent). H is the height above the Mean Sea Level. The geopotential height can be understood as a scaled geopotential, similar to the *dynamic height* in geodesy. Figure A.1 shows that the differences between the geometric height and geopotential height as a function of the geometric height at different latitudes. The differences increase with the altitude. We now introduce an approximate method to convert between the geometric (strictly speaking, the height should be the orthometric height) and the geopotential height without using complicated gravity models.

If we approximate the gravity at H by applying the inverse square law:

$$g(H,\phi) = g(\phi)\left(\frac{R}{R+H}\right)^2 \tag{A.2}$$

Here R is an appropriate radius of the Earth at a given geodetic latitude. $g(\phi)$ is the gravity on the geoid at the latitude ϕ. H is the height of the computational point above the geoid. Substituting (A.2) into (A.1), we obtain

$$\Phi = \frac{g(\phi)}{g_0}\frac{RH}{R+H} \tag{A.3}$$

H can be derived as:

$$H = \frac{R\Phi}{\frac{g(\phi)}{g_0}R - \Phi} \tag{A.4}$$

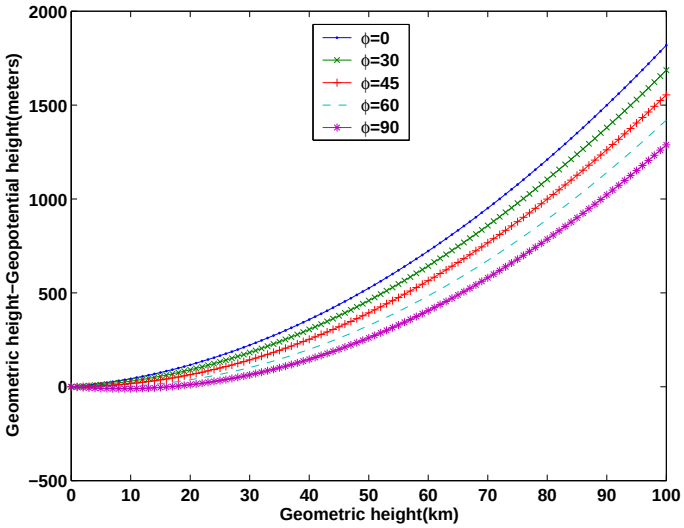


Figure A.1: The geometric height minus geopotential height derived using the EGM 96 model.

In reality, Earth is not a non-rotating, homogeneous sphere, its gravity does not strictly follow the inverse square law. Meanwhile, as the Earth is rotating, the centrifugal component of the gravity tends to increase with the increase of the parallel radius. To compensate the effect of the gravity variation and centrifugal force but still keep the form of expression in (A.3) and (A.4), a strategy suggested by W.D. Lambert[*List*, 1951] is to adjust R as a fictitious radius for the expressions to obtain the right value. Differentiating the equation (A.2) with respect to H and evaluating at $H = 0$, we obtain

$$R_f(\phi) = \frac{2g(\phi)}{-\frac{\partial g(H,\phi)}{\partial H}\Big|_{H=0}} \tag{A.5}$$

R_f is the fictitious radius and is used to replace R in (A.3) and (A.4).

The gravity at the geoid can be approximated by the *normal gravity*. From *Heiskanen and Moritz* [1967], the normal gravity above the reference ellipsoid can be computed using equation (2-123):

$$g(H,\phi) \approx \gamma_H = \gamma(\phi)\left[1 - \frac{2}{a}(1 + f + m - 2f\sin^2\phi)H + 3\left(\frac{H}{a}\right)^2\right] \tag{A.6}$$

Following quation (A.5), yields,

$$R_f(\phi) = \frac{a}{1 + f + m - 2f\sin^2\phi} \tag{A.7}$$

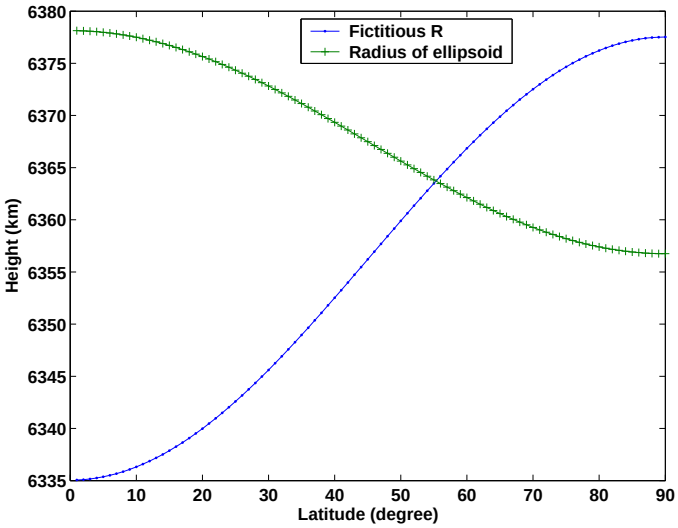


Figure A.2: The fictitious radius and the radius of reference ellipsoid.

The variables can be replaced from *Moritz* [1992], finally, (A.3) and (A.4) becomes

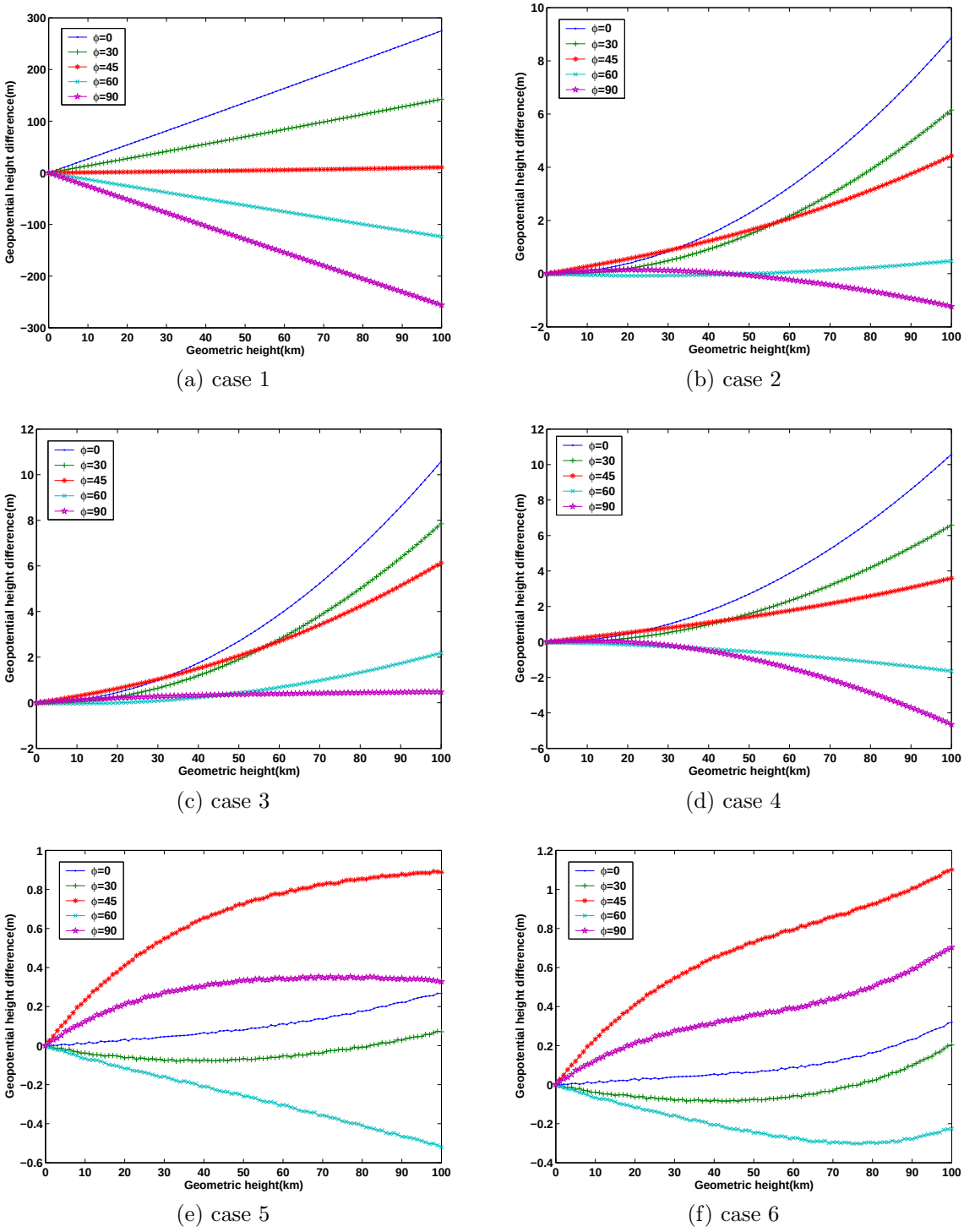
$$\Phi = \frac{g(\phi)}{g_0}\frac{R_f H}{R_f + H} \tag{A.8}$$

$$H(\Phi, \phi) = \frac{R_f(\phi)\Phi}{\frac{g(\phi)}{g_0}R_f(\phi) - \Phi} \tag{A.9}$$

In the literature, there exists simplified versions of above formulas. In order to verify the accuracy of these formulas. We compare the geopotential height calculated using EGM 96 and the following variations of formulas:

1. $\Phi = \frac{aH}{a+H}$, a = 6,378.137 km
2. $\Phi = \frac{g(\phi)}{g_0}\frac{aH}{a+H}$, a = 6,371 km
3. $\Phi = \frac{g(\phi)}{g_0}\frac{aH}{a+H}$, a = 6,378.137 km
4. $\Phi = \frac{g(\phi)}{g_0}\frac{aH}{a+H}$, a = radius of the ellipse (7.8)
5. $\Phi = \frac{g(\phi)}{g_0}\frac{R_f H}{R_f+H}$ (A.8)
6. Numerical integration of the equation (A.6)

The differences between each of these six cases and the EGM 96 derived geopotential height are shown in Figure A.3. Considering the latitudinal variation of gravity g greatly reduces the conversion error. Case 1 gives the worst result, which is the formula used in the GRACE AOD document [*Flechtner*, 2005]. After further adopting the Smithsonian radius R_f, the error was reduced by one order of magnitude. The numerical integration formula performs equally well with equation (A.8), but it is less favorable since equation (A.8) is very simple and easy to implement.



(a) case 1

(b) case 2

(c) case 3

(d) case 4

(e) case 5

(f) case 6

Figure A.3: Different formulas to compute geopotential height.

APPENDIX B

THE COMPUTATION OF SURFACE PRESSURE

To account for the fact that the artificial topography in numerical models are different from the realistic Earth surface, it is needed to compute the surface pressure from upper air analyses. In a large scale, the atmosphere is assumed to be in static equilibrium in the vertical direction, which is the hydrostatic equilibrium (2.35):

$$dP = -g(z)\rho(z)dz \tag{B.1}$$

z is the geometric height. In meteorological products, geopotential heights are commonly used. The equation can also be written in terms of the geopotential height from the definition of equation (A.1):

$$dP = -g_0\rho(\Phi)d\Phi \tag{B.2}$$

Φ represents the geopotential height. The density ρ in the above equation can be derived from the ideal gas law (or equation of state) (7.10):

$$P = \rho R T_v \qquad \text{or} \qquad \rho = \frac{P}{RT_v} \tag{B.3}$$

Substituting (B.3) into (B.2), yields

$$\ln\frac{P_2}{P_1} = -\frac{g_0}{RT_v}(\Phi_2 - \Phi_1) \tag{B.4}$$

See equation (7.10) for the definition of the variables.

If the topographic surface is between two mandatory pressure levels (smaller than P_1 and larger than P_2), the surface pressure can be derived as:

$$\ln\frac{P_s}{P_1} = \ln\frac{P_2}{P_1}\frac{\Phi_s - \Phi_1}{\Phi_2 - \Phi_1} \tag{B.5}$$

Where P_s is the surface pressure. The above equation assumes the virtual temperature T_v is a constant in the layer between P_1 and P_2.

Sometimes, the topographic surface is below the lowest mandatory pressure level P_1. If we can assume a uniform temperature lapse rate $\alpha = 6.5$ K/km, the surface temperature can be extended as:

$$T_s = T_1 + \alpha(\Phi_s - \Phi_1) \tag{B.6}$$

The surface pressure is derived from (B.4):

$$P_s = P_1 \exp\left(- \frac{g_0(\Phi_s - \Phi_1)}{R \cdot 0.5(T_s + T_1)} \right) \tag{B.7}$$

For the case that only sea level pressure P_0 and temperature T_0 are available, the surface pressure can be obtained in a similar way. The surface temperature is:

$$T_s = T_0 - \alpha\Phi_s \tag{B.8}$$

and surface pressure

$$P_s = P_0 \exp\left(- \frac{g_0\Phi_s}{R \cdot 0.5(T_s + T_0)} \right) \tag{B.9}$$

APPENDIX C

THE LIST OF AUTOMATIC WEATHER STATIONS

ID	Station	Lat	Lon	Height (m)	GPS Surveyed
		Adelie Coast			
08909	Port Martin	66.820 S	141.400 E	39	0
08914	D-10	66.710 S	139.830 E	243	0
08986	D-47	67.397 S	138.726 E	1550	0
08988	Cape Denison	67.009 S	142.664 E	31	0
08989	Dome C II	75.121 S	123.374 E	3250	0
		High Polar Plateau			
08904	Dome Fuji	77.310 S	39.700 E	3810	0
08918	Relay Station	74.017 S	43.062 E	3353	0
08985	Henry	89.011 S	1.025 W	2755	0
21356	Clean Air	90.000 S	0.000 E	2835	0
21359	Mizuho	70.700 S	44.290 E	2260	0
		West Antarctica			
08900	Harry	83.003 S	121.393 W	945	0
08903	Byrd	80.007 S	119.404 W	1530	0
08931	Brianna	83.889 S	134.154 W	525	1
08938	Siple Dome	81.656 S	148.773 W	668	1
08981	Mount Siple	73.198 S	127.052 W	230	0
21358	Theresa	84.599 S	115.811 W	1463	0

Continued

Table C.1: The automatic weather stations in Antarctica.

Table C.1 Continued

ID	Station	Lat	Lon	Height (m)	GPS Surveyed
		Ross and Ocean Island (McMurdo Area)			
08695	Cape Spencer	77.967 S	167.531 E	24	1
08697	Herbie Alley	78.100 S	166.670 E	24	0
08901	Cape Bird	77.224 S	166.440 E	42	1
08906	Marble Point	77.439 S	163.754 E	108	1
08927	Windless Bight	77.728 S	167.703 E	60	0
08929	Ferrell	77.884 S	170.818 E	45	1
08937	Pegasus South	77.990 S	166.568 E	5	1
21357	Pegasus North	77.952 S	166.500 E	8	1
21360	Laurie II	77.529 S	170.807 E	38	1
21362	Linda	78.451 S	168.394 E	43	1
21364	Willie Field	77.866 S	166.983 E	14	1
		Ross Ice Shelf			
08905	Manuela	74.946 S	163.687 E	78	0
08907	Whitlock	76.144 S	168.392 E	275	0
08984	Possession Island	71.891 S	171.210 E	30	0
08908	Lettau	82.518 S	174.452 W	30	0
08911	Gill	79.985 S	178.611 W	25	0
08913	Schwerdtfeger	79.875 S	170.105 E	54	1
08915	Elaine	83.134 S	174.169 E	60	0
08934	Marilyn	79.954 S	165.130 E	75	0
		Weddell Sea/Peninsula Region			
08917	Sky Blu	74.792 S	71.488 W	1395	0
08920	Uranus Glacier	71.430 S	68.930 W	780	0
08923	Bonaparte Point	64.778 S	64.067 W	8	0
08926	Larsen Ice Shelf	66.949 S	60.897 W	17	0
08947	Racer Rock	64.067 S	61.613 W	17	0

APPENDIX D

SPHERICAL HARMONIC ANALYSIS WITH BLOCK MEAN

The equations to convert the 3-D meteorological fields to spherical harmonic coefficients have the common form of:

$$\begin{bmatrix} C_{nm} \\ S_{nm} \end{bmatrix} = \frac{a^2}{(2n+1)M} \int_0^{2\pi} \int_0^{\pi} f(\theta,\lambda,n) P_{nm}(\cos\theta) \begin{bmatrix} \cos m\lambda \\ \sin m\lambda \end{bmatrix} \sin\theta d\theta d\lambda \tag{D.1}$$

where $f(\theta,\lambda,n)$ represents the vertical integral in equations (7.11) or (7.17). The value of f was evaluated as an area mean using the four corner points of the equiangular blocks on the reference ellipsoid. There are $N = \pi/\Delta\lambda$ blocks in the latitude direction and $2N$ blocks in the longitude direction. A numerical quadrature formula can be formed using a digitized version of (D.1) as suggested by *Colombo* [1981]:

$$\begin{bmatrix} C_{nm} \\ S_{nm} \end{bmatrix} = \frac{a^2}{(2n+1)M} \sum_{i=0}^{N-1} \sum_{j=0}^{2N-1} \bar{f}_{i,j}(n) \int_{\theta_i}^{\theta_{i+1}} P_{nm}(\cos\theta)\sin\theta d\theta \int_{\lambda_j}^{\lambda_{j+1}} \begin{bmatrix} \cos m\lambda \\ \sin m\lambda \end{bmatrix} d\lambda \tag{D.2}$$

The above formula can be re-written as:

$$\begin{bmatrix} C_{nm} \\ S_{nm} \end{bmatrix} = \frac{a^2}{(2n+1)M} \sum_{i=0}^{N-1} \sum_{j=0}^{2N-1} \bar{f}_{i,j}(n)\ IP^i_{nm} \begin{bmatrix} IC^j_m \\ IS^j_m \end{bmatrix} \tag{D.3}$$

where

$$IP^i_{nm} = \int_{\theta_i}^{\theta_{i+1}} P_{nm}(\cos\theta)\sin\theta d\theta \tag{D.4}$$

and

$$\begin{bmatrix} IC^j_m \\ IS^j_m \end{bmatrix} = \int_{\lambda_j}^{\lambda_{j+1}} \begin{bmatrix} \cos m\lambda \\ \sin m\lambda \end{bmatrix} d\lambda \tag{D.5}$$

The integral in equation (D.5) has a closed analytical solution along the longitude. We will now give a recursive formula to compute the integrals of the associated Legendre functions in (D.4).

The fully normalized associated Legendre functions can be computed by the following recursive formulas [*Heiskanen and Moritz*, 1967; *Jekeli*, 1996]:

$$P_{n,n}(\cos\theta) = W_{n,n}\sin\theta P_{n-1,n-1}(\cos\theta) \tag{D.6}$$

$$P_{n,n-1}(\cos\theta) = W_{n,n}\cos\theta P_{n-1,n-1}(\cos\theta) \tag{D.7}$$

$$P_{n,m} = W_{n,m}\Big(\cos\theta P_{n-1,m}(\cos\theta) - W^{-1}_{n-1,m}P_{n-2,m}(\cos\theta)\Big) \tag{D.8}$$

where $m = 0, 1, 2, \cdots$, $n \geq m$, and

$$W_{1,1} = \sqrt{3} \tag{D.9}$$

$$W_{n,n} = \sqrt{\frac{2n+1}{2n}} \qquad \text{for} \qquad n > 1 \tag{D.10}$$

$$W_{n,m} = \sqrt{\frac{(2n+1)(2n-1)}{(n+m)(n-m)}} \tag{D.11}$$

These recursive relations can be initialized by $P_{0,0} = 1$. Based on the above formulas, one can derive the following recursive formulas to compute the integrals of the fully normalized associated Legendre function (D.4) [*Paul*, 1978]:

$$IP_{n,n} = \frac{W_{n,n}}{n+1}\Big[nW_{n-1,n-1}IP_{n-1,n-2} - \cos\theta_2 W^{-1}_{n,n}P_{n,n}(\cos\theta_2) + \cos\theta_1 W^{-1}_{n,n}P_{n,n}(\cos\theta_1)\Big] \tag{D.12}$$

$$IP_{n,n-1} = \frac{W_{n,m}}{m+1}\Big[\sin^2\theta_2 P_{n-1,m}(\cos\theta_2) - \sin^2\theta_1 P_{n-1,m}(\cos\theta_1)\Big] \tag{D.13}$$

$$IP_{n,m} = \frac{W_{n,m}}{m+1}\Big[(n-2)W^{-1}_{n-1,m}IP_{n-2,m} + \sin^2\theta_2 P_{n-1,m}(\cos\theta_2) - \sin^2\theta_1 P_{n-1,m}(\cos\theta_1)\Big] \tag{D.14}$$

The above recursive relations are initialized by:

$$IP_{0,0} = \cos\theta_1 - \cos\theta_2 \tag{D.15}$$

$$IP_{1,1} = \frac{W_{1,1}}{2}\Big[\theta_2 - \theta_1 - \cos\theta\sin\theta\Big|_{\theta 1}^{\theta_2}\Big] \tag{D.16}$$

BIBLIOGRAPHY

Anthes, R. A., C. Rocken, and Y. H. Kuo (2000), Applications of COSMIC to meteorology and climate, *Terr. Atmos. Oceanic Sci.*, *11*, 115–156.

Arnold, V. I. (1978), *Mathematical methods of classical mechanics*, Springer-Verlag, New York.

Bettadpur, S. (2004), *GRACE UTCSR level-2 processing standards documents RL0001*, Center for Space Research, The University of Texas at Austin.

Bevis, M., S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. Ware (1992), GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system, *J. Geophys. Res.*, *97*, 15,787–15,801.

Beyerle, G., T. Schmidt, G. Michalak, S. Heise, J. Wickert, and C. Reigber (2005), GPS radio occultation with GRACE: Atmospheric profiling utilizing the zero difference technique, *Geophys. Res. Lett.*, *32*, L13806, doi:10.1029/2005GL023109.

Blewitt, G., D. Lavallee, P. Clarke, and K. Nurutdinov (2001), A new global mode of Earth deformation: Seasonal cycle detected, *Science*, *294*, 2342–2345.

Born, M., and E. Wolf (1993), *Principles of Optics*, Pergamon Press, Oxford, UK.

Bouttier, F., and P. Courtier (1999), Data assimilation concepts and methods, Lecture Notes, ECMWF.

Boy, J.-P., and B. F. Chao (2005), Precise evalutation of atmospheric loading effects on Earth's time-variable gravity field, *J. Geophys. Res.*, *110*, B08412, doi:10.1029/2002JB002333.

Boy, J.-P., P. Gegout, and J. Hinderer (2002), Reduction of surface gravity data from global atmospheric pressure loading, *Geophys. J. Int.*, *149*, 534–545.

Bromwich, D. H., and R. L. Fogt (2004), Strong trends in the skill of the ERA-40 and NCEP/NCAR Reanalyses in the high and middle latitudes of the southern hemisphere, 1958-2001, *J. Climate*, *17*, 4603–4619.

Bromwich, D. H., R. I. Cullather, and R. W. Grumbine (1999), An assessment of the NCEP operational global spectral model forecasts and analyses for Antarctica during FROST, *Weather Forecast.*, *14*, 835–850.

Bromwich, D. H., J. J. Cassano, T. Klein, G. Heinemann, K. M. Hines, K. Steffen, and J. E. Box (2001), Mesoscale modeling of katabatic winds over Greenland with the polar MM5, *Mon. Weather Rev.*, *129*, 2290–2309.

Chambers, D. P., J. Wahr, and R. S. Nerem (2004), Preliminary observations of global ocean mass variations with GRACE, *Geophys. Res. Lett.*, *31*, L13310, doi:10.1029/2004GL020461.

Chao, B. F., and A. Y. Au (1991), Temporal variation of the Earth's low-degree zonal gravity field caused by atmospheric mass redistribution: 1980-1988, *J. Geophys. Res.*, *96*(B4), 6569–6575.

Chao, B. F., W. P. O'Connor, A. T. C. Chang, D. K. Hall, and J. L. Foster (1987), Snow load effect on the Earth's rotation and gravitational field, 1979-1985, *J. Geophys. Res.*, *92*(B9), 9415–9422.

Chapman, S., and R. S. Lindzen (1970), *Atmospheric tides: Thermal and gravitational*, Gordon and greach, Science Publishers, Inc., New York.

Chen, J. L., C. R. Wilson, R. J. Eanes, and R. S. Nerem (1999), Geophysical interpretation of observed geocenter variations, *J. Geophys. Res.*, *104*(B2), 2683–2690.

Chen, J. L., M. Rodell, C. R. Wilson, and J. S. Famiglietti (2005a), Low degree spherical harmonic influences on Gravity Recovery and Climate Experiment (GRACE) water storage estimates, *Geophys. Res. Lett.*, *32*, L14405, doi:10.1029/2005GL022964.

Chen, J. L., C. R. Wilson, J. S. Famiglietti, and M. Rodell (2005b), Spatial sensitivity of the Grace Recovery and Climate Experiment (GRACE) time-variable gravity observations, *J. Geophys. Res.*, *110*, B08408, doi:10.1029/2004JB003536.

Cheng, M. K., and B. D. Tapley (1999), Seasonal variations in low degree zonal harmonics of the Earth's gravity field from satellite laser ranging observations, *J. Geophys. Res.*, *104*(B2), 2667–2681.

Chew, W. C. (1990), *Waves and fields in inhomogeneous media*, Van Nostrand Reinhold, New York.

Colombo, O. L. (1981), Numerical methods for harmonic analysis on the sphere, *OSU Report 310*, Dept. of Geod. Sci. and Surv., The Ohio State University, Columbus, Ohio.

Cox, C. M., and B. F. Chao (2002), Detection of a large-scale mass redistribution in the terrestial system since 1998, *Science*, *297*, 831–833.

Cullather, R. I., D. H. Bromwich, and R. W. Grumbine (1997), Validation of operational numerical analysis in Antarctic latitudes, *J. Geophys. Res.*, *102*(D12), 13,761–13,784.

Davis, C. H., Y. Li, J. R. McConnell, M. M. Frey, and E. Hanna (2005), Snow-fall-driven growth in east Antarctic ice sheet mitigates recent sea-level rise, *Science*, *308*(5730), 1898–1901, doi:10.1126/science.1110662.

Dickey, J. O., S. L. Marcus, O. de Viron, and i. Fukumori (2002), Recent Earth oblateness variations: Unveiling climate and postglacial rebound, *Science*, *298*, 1975–1977.

Egorov, Y. V., A. I. Komech, and M. Shubin (1999), *Elements of the Modern theory of partial differential equations*, Springer-Verlag, New York.

Eyre, J. R. (1994), Assimilation of radio occultation measurements into a numerical weather prediction system, *Technical Memorandum 199*, ECMWF.

Farrel, W. (1972), Deformation of the Earth by surface loads, *Review of Geophysics*, *10*, 761–797.

Fjeldbo, G., and V. R. Eshleman (1969), Atmosphere of Venus as studied with the Mariner 5 dual readio-frequency occultation experiment, *Radio Sci.*, *4*, 879–897.

Fjeldbo, G., A. J. Kliore, and V. R. Eshleman (1971), The neutral atmosphere of Venus as studied with the Mariner V radio occultation experiments, *The Astronomical Journal*, *76*(2), 123–140.

Flechtner, F. (2005), *AOD1B Product description Document*, GeoForschungsZentrum Postdam, Postdam, Germany, 2.1 ed.

Fleming, E. L., M. R. Chandra, M. R. Shoeberl, and J. J. Barnett (1988), Monthly mean global climatology of temperature, wind, geopotential height and pressure for 0-120 km, *Technical Memorandum 100697*, National Aeronautics and Space Administration, Washinton, D. C.

Fu, L.-L., and G. Pihos (1994), Determining the response of sea level to atmospheric pressure forcing using TOPEX/POSEIDON data, *J. Geophys. Res.*, *99*, 24,633–24,646.

Ge, S., C. Shum, J. Wickert, and C. Reigber (2002), GPS occultation: Potential a new data source for improvement of antarctic pressure field, *Wuhan University Journal of natural science*, *8*(28), 636–648.

Ge, S., C. K. Shum, S.-C. Han, and D. H. Bromwich (2005), Assessment of ECMWF and polar MM5 over Antarctica using GRACE, in *IAMAS 2005*, Beijing, China.

Gorbunov, M. E. (2001), Radioholographic methods for processing radio occultation data in multipath regions, *Scientific Report 01-02*, Danish Meteorological Institute, Copenhagen.

Gorbunov, M. E. (2002), Canonical transform method for processing radio occultation data in the lower troposphere, *Radio Sci.*, *37*(5), 1076, doi:10.1029/2000RS002592.

Gorbunov, M. E., and S. V. Sokolovsky (1993), Remote sensing of refractivity from space for global observations of atmospheric parameters, *Report 119*, Max Planck Institute for Meteorology, Hamburg, Germany.

Gorbunov, M. E., S. V. Sokolovsky, and L. Bengtsson (1996), Space refractive tomography of the atmosphere: modeling of direct and inverse problems, *Report 210*, Max Planck Institute for Meteorology, Hamburg, Germany.

Gorbunov, M. E., A. S. Gurvich, and L. Kornblueh (2000), Comparative analysis of radioholographic methods of processing radio occultation data, *Radio Sci.*, *35*(4), 1025–1034.

Gordon, B. A. (2002), *The WMO format for the storage of weather product information and the exchange of weather product messages in gridded binary form as used by NCEP central operations*, NCEP Central Operations, 1 ed., office note 388.

Gruber, J., Th. Wunsch, S. Bettadpur, and T. Pekker (2001), Atmospheric and oceanic gravity field de-aliasing for GRACE science data processing, in *Geophysical research abstracts, 26th General Assembly*, vol. 3, Nice, France.

Han, S., C. Jekeli, and C. K. Shum (2004), Time-variable aliasing effects of ocean tides, atmosphere and continental water mass on monthly mean GRACE gravity field, *J. Geophys. Res.*, *109*, B04403, doi:10.1029/2003JB002501.

Healy, S., and J. Eyre (2000), Retrieving temperature, water vapor and surface pressure information from refractivity-index profiles derived by radio occultation: A simulation study, *Quart. J. Roy. Meteorol. Soc.*, *126*, 1661–1683.

Healy, S. B. (2001), Smoothing radio occultation bending angles above 40 km, *Ann. Geophys.*, *19*, 459–468.

Healy, S. B., A. M. Jupp, and C. Marquardt (2005), Forecast impact experiment with GPS radio occultation measurements, *Geophys. Res. Lett.*, *32*, L03804, doi:10.1029/2004GL020806.

Hedin, A. E. (1991), Extension of the msis thermosphere model into the middle and lower atmosphere, *J. Geophys. Res.*, *96*(A2), 1159–1172.

Heiskanen, W., and H. Moritz (1967), *Physical Geodesy*, W. H. Freeman Publications Co., San Francisco.

Hines, K. M., D. H. Bromwich, and G. J. Marshall (2000), Artificial surface pressure trends in the NCEP-NCAR reanalysis over the routhern ocean and Antarctica, *J. Climate*, *13*, 3,940–3,952.

Hocke, K. (1997), Inversion of GPS meteorology data, *Ann. Geophys.*, *15*, 443–450.

Hocke, K., A. Pavelyev, O. Yakovlev, L. Barthes, and N. Jakowski (1999), Radio occultation data analysis by the radio holographic method, *J. Atmos. Solar-Terr. Phys.*, *61*, 1169–1177.

Holton, J. R. (1992), *An introduction to dynamic meteorology*, *International Geophysics series*, vol. 48, third ed., Academic Press, San Diego.

Igarashi, K., A. Pavelyev, D. Hocke, K. Pavelyev, I. A. Kucherjavenkov, S. Matyugov, A. Zakharov, and O. Yakovlev (2000), Radio holographic principle for observing natural processes in the atmosphere and retrieving meteorological parameters from radio occultation data, *Earth, Planets, and Space*, *52*(11), 893–899.

Jayne, S. R., and J. M. Wahr (2003), Observing ocean heat content using satellite gravity and altimetry, *J. Geophys. Res.*, *108*(C2), 3031, doi:10.1029/2002JC001619.

Jekeli, C. (1981), Alternative methods to smooth the Earth's gravity field, *Tech. Rep. 327*, Dept. of Geodetic Science and Surveying, Columbus, Ohio.

Jekeli, C. (1996), Spherical harmonic analysis, aliasing, and filtering, *Journal of Geodesy*, *70*(214-223).

Jekeli, C. (1999), The determination of gravitational potential differences from satellite-to-satellite tracking, *Clestial Mechanics and Dynamic Astronomy*, *75*, 85–101.

Jekeli, C. (2000), Geometric geodesy, Class Notes, Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University.

Jensen, A. S., M. S. Lohmann, H.-H. Benzon, and A. S. Nielsen (2003), Full spectrum inversion of radio occultation signals, *Radio Sci.*, *38*(3), 1040, doi:10.1029/2002RS002763.

Jensen, A. S., M. S. Lohmann, A. S. Nielsen, and H.-H. Benzon (2004), Geometrical optics phase matching of radio occultation signals, *Radio Sci.*, *39*, RS3009, doi:10.1029/2003RS002899.

King, J. C., and J. Turner (1997), *Antarctic Meteorology and Climatology*, Cambridge University Press, Cambridge, U. K.

Kravtsov, Y. A., and Y. I. Orlov (1990), *Geometrical optics of inhomogeneous media*, Springer-Verlag, New York.

Kuo, Y.-H., W. S. Schreiner, J. Wang, D. L. Rossiter, and Y. Zhang (2005), Comparison of GPS radio occultation soundings with radiosondes, *Geophys. Res. Lett.*, *32*, L05817, doi:10.1029/2004GL021443.

Kursinski, E. R., G. A. Hajj, W. I. Bertiger, S. S. Leroy, T. K. Meehan, L. J. Romans, J. T. Schofield, D. J. McCleese, W. G. Melbourne, C. L. Thornton, T. P. Yunck, J. R. Eyre, and R. N. Nagatani (1996), Initial results of radio occultation observations of Earth's atmosphere using the Global Positioning System, *Science*, *271*, 1,107–1,110.

Leroy, S. S. (1997), Measurement of geopotential heights by GPS radio occultation, *J. Geophys. Res.*, *102*(D6), 6971–6986.

Lindal, G., J. Lyons, D. Sweetnam, V. Eshleman, D. P. Hinson, and G. L. Tyler (1987), The atmosphere of Uranus: Results of radio occultation measurements with Voyager 2, *J. Geophys. Res.*, *92*(13), 14,987–15,001.

List, R. J. (1951), *Smithsonian meteorological tables*, *Smithsonian Mescellaneous Collections*, vol. 114, sixth ed., The Smithsonian Institution, City of Washingtion.

Liston, G. E., and J.-G. Winther (2005), Antarcitc surface and subsurface snow and ice melt fluxes, *J. Climate*, *18*(1469-1481).

Liu, H., X. Zou, H. Shao, J. C. Anthes, R. A.and Chang, J.-H. Tsing, and H. Wang (2001), Impact of 837 GPS/MET bending angle profiles on assimilation and forecasets for the period June 20-30, 1995, *J. Geophys. Res.*, *106*(D23), 31,771–31,786.

Marquardt, C., K. Labitzke, C. Reigber, T. Schmidt, and J. Wickert (2001), An assessment of the quality of GPS/MET radio limb soundings during february 1997, *Phys. Chem. Earth*, *26*(3), 125–130.

Marshall, G. J. (2002), Trends in antarctic geopotential height and temperature: A comparison between radiosonde and NCEP-NCAR reanalysis data, *J. Climate*, *15*(659-674).

Martin, J. (1992), *Wave Propagation in Random Media (Scintillations)*, chap. Simulation of wave propagation in random media: Theory and application, pp. 463–486, SPIE-The Int. Soc. for Opt. Eng., Bellingham, Wash.

Melbourne, W. G., E. S. Davis, C. B. Duncan, G. A. Hajj, K. R. Hardy, T. K. Kursinski, T. K. Meehan, L. E. Young, and T. P. Yunck (1994), The applicaiton of spacebourne GPS to atmospheric limb sounding and global change mornitoring, *JPL Publication 94-18*, Jet Propulsion Lab, Pasadena, CA, 147pp.

Mishchenko, A. S., V. E. Shatalov, and B. Y. Sternin (1990), *Lagrangian Manifolds and the Maslov Operator*, Springer-Verlag, New York.

Moritz, H. (1992), Geodetic reference system 1980, *Bull. Geod.*, *66*(2), 187–192.

Nerem, R. S., R. J. Eanes, P. F. Thompson, and J. L. Chen (2000), Observations of annual variations of the Earth's gravitational field using satellite laser ranging and geophysical models, *Geophys. Res. Lett.*, *27*(12), 1783–1786.

Nerem, R. S., J. M. Wahr, and E. W. Leuliette (2003), Measureing the distribution of ocean mass using GRACE, *Space Sci. Rev.*, *108*, 331–344.

Parish, T. R., and D. H. Bromwich (1998), A case study of antarctic katabatic wind interaction with large-scale forcing, *Mon. Weather Rev.*, *126*, 199–209.

Paul, M. K. (1978), Recurrence relations for integrals of the associated legendre functions, *Bulletin Géodésique*, *52*, 177–190.

Pendlebury, S. F., N. D. Adams, T. L. Hart, and J. Turner (2003), Numerical weather prediction model performance over high southern latitudes, *Mon. Weather Rev.*, *131*, 335–353.

Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Arkin (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issures, *J. Geophys. Res.*, *107*(A12), 1468, doi:10.1029/2002JA009430.

Poli, P., C. O. Ao, M. de la Torre Jurez, J. Joiner, G. A. Hajj, and R. M. Hoff (2003), Evaluation of CHAMP radio occultation refractivity using data assimilation office analyses and radiosondes, *Geophys. Res. Lett.*, *30*(15), 1800, doi:10.1029/2003GL017637.

Ponte, R. M., and J. Dorandeu (2003), Uncertainties in ECMWF surace pressure fields over the ocean in relation to sea level analysis and modeling, *J. Atmos. Ocean. Technol.*, *20*, 301–307.

Ponte, R. M., and P. Gaspar (1999), Regional analysis of the invert barometer effect over the global ocean using TOPEX/POSEIDON data and model results, *J. Geophys. Res.*, *104*(C7), 15,587–15,601.

Ponte, R. M., and R. D. Ray (2002), Atmospheric pressure corrections in geodesy and oceanography: A strategy for handling air tides, *Geophys. Res. Lett.*, *29*(24), 2153, doi:10.1029/2002GL016340.

Rannon, P. R., C. H. Bishop, and J. B. Kerr (1997), Does the surface pressure equal the weight per unit area of a hydrostatic atmosphere, *Bull. Amer. Meteorol. Soc.*, *78*(11), 2637–2642.

Ray, R. D. (2001), Comparisons of global analyses and station observations of the S_2 barometric tide, *J. Atmos. Solar-Terr. Phys.*, *63*, 1085–1097.

Ray, R. D., and R. M. Ponte (2003), Barometric tides from ECMWF operational analyses, *Ann. Geophys.*, *21*, 1897–1910.

Reigber, C., H. Luehr, and P. Schwintzer (2002), CHAMP mission status, *Adv. Space Res.*, *30*(2), 129–134.

Rignot, E., and R. H. Thomas (2002), Mass balance of polar ice sheets, *Science*, *297*(5586), 1502–1506, doi:10.1126/science.1073888.

Rocken, C., R. Anthes, M. Exner, D. Hunt, S. Sokolovskiy, R. Ware, M. Gorbunov, W. Schreiner, D. Feng, B. Herman, Y.-H. Kuo, and X. Zou (1997), Analysis and validation of GPS/MET data in the neutral atmosphere, *J. Geophys. Res.*, *102*, 29,849–29,866, doi:10.1029/97JD02400.

Rodell, M., and J. S. Famiglietti (1999), Detectability of variations in continental water storage from satellite observations of the time dependent gravity field, *J. Geophys. Res.*, *35*(9), 2705–2723.

Rodell, M., and J. S. Famiglietti (2001), An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, *37*(5), 1327–1339.

Rodell, M., J. Famiglietti, J. Chen, S. I. Seneviratne, P. Viterbo, S. Holl, and C. R. Wilson (2004a), Bason scale estimates of evapotranspiration using GRACE and other observtions, *Geophys. Res. Lett.*, *31*, L20504, doi:10.1029/2004GL020873.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll (2004b), The global land data assimilation system, *Bull. Amer. Meteorol. Soc.*, *85*(3), 381–394.

Simmonds, I., and D. A. Jones (1998), The mean structure and temporal variability of the semiannual oscillation in the soughtern extratropics, *Int. J. Climatol.*, *18*(473-504).

Smith, E. K., and S. Weintraub (1953), The constants in the equation of atmospheric refractive index at radio frequencies, in *Proceedings of the Institute of Radio Engineers*, vol. 41, pp. 1035–1037.

Sokolovskiy, S. (2004), Open loop tracking and interting GPS radio occultation signals: Simulation study, in *Occultations for probing atmosphere and climte*, edited by G. Kirchengast, U. Foelsche, and A. K. Steiner, pp. 39–51, Springer-Verlag, Berlin.

Sokolovskiy, S. V. (2000), Inversions of radio occultation amplitude data, *Radio Sci.*, *35*(1), 97–105.

Sokolovskiy, S. V. (2001), Modeling and inverting radio occultation signals in the moist troposphere, *Radio Sci.*, *36*(3), 441–458.

Sokolovskiy, S. V., and D. Hunt (1996), Statistical optimization approach for GPS/MET data inversions, in *Proceedings of URSI GPS/MET Workshop*, Tucson, Arizona.

Swenson, S., and J. Wahr (2002), Estimated effects of the vertical structure of atmospheric mass on the time variable geoid, *J. Geophys. Res.*, *107*(B9), 2194, doi:10.1029/2000JB000024.

Swenson, S., J. Wahr, and P. C. D. Milly (2003), Estimated accuracies of regional water storage variations inferred from the Gravity Recovery and Climate Experiment (GRACE), *Water Resour. Res.*, *39*(8), 1223, doi:10.1029/2002WR001808.

Swenson, S. C., and P. C. D. Milly (2006), Climate model biases in seasonality of continental water storage revealed by satellite gravimetry, *Water Resour. Res.*, *42*, W03201, doi:10.1029/2005WR004628.

Syndergaard, S. (1998), Modeling the impact of the Earth's oblateness on the retrieval of temperature and pressure profiles from limb sounding, *J. Atmos. Solar-Terr. Phys.*, *60*(2), 171–180.

Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins (2004a), GRACE measurements of mass variability in the Earth system, *Science*, *305*, 503–505.

Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber (2004b), The Gravity Recovery and Climate Experiment: Mission overview and early results, *Geophys. Res. Lett.*, *31*, L09607, doi:10.1029/2004GL019920.

Tapley, B. D., J. Ries, S. Bettadpur, D. Chambers, M. Cheng, F. Condi, B. Gunter, Z. Kang, P. Nagel, R. Pastor, T. Pekker, S. Poole, and F. Wang (2005), GGM02- an improved earth gravity field model from GRACE, *Journal of Geodesy*, *79*(8), 467–478, doi:10.1007/s00190-005-0480-z.

Thompson, P. F., S. V. Bettadpur, and B. D. Tapley (2004), Impact of short period, non-tidal, temporal mass variability on GRACE gravity estimates, *Geophys. Res. Lett.*, *31*, L06619, doi:10.1029/2003GL019285.

Trenberth, K. E. (1981), Seasonal variations in global sea level pressure and the total mass of the atmosphere, *J. Geophys. Res.*, *86*(C6), 5238–5246.

Trenberth, K. E. (1992), Global analyses from ECMWF and atlas of 1000 to 10 mb circulation statistics, *NCAR Tech. Note NCAR/TN-373+STR*, National Center for Atsmospheric Research, Boulder, Colorado.

Trenberth, K. E., and C. J. Guillemot (1994), The total mass of the atmosphere, *J. Geophys. Res.*, *99*(D11), 23,079–23,088.

Trenberth, K. E., and L. Smith (2005), The mass of the atmosphere: A constraint on global analyses, *J. Climate*, *18*, 864–875.

U. S. Standard Atmosphere (1976), U.S standard atmosphere 1976, *Tech. rep.*, U.S. Government Printing Office, Washington, D. C.

van Dam, T. M., and J. Wahr (1987), Dispalcements of the Earth's surface due to atmospheric loading effects on gravity and baseline measurements, *J. Geophys. Res.*, *99*(B12), 1,281–1,286.

van Dam, T. M., and J. Wahr (1993), The atmospheric load response of the ocean determined using Geosat altimeter data, *Geophys. J. Int.*, *113*, 1–16.

van Dam, T. M., G. Blewitt, and M. B. Heflin (1994), Atmospheric pressure loading effects on Global Positioning System coordinate determinations, *J. Geophys. Res.*, *99*(B12), 23,939–23,950.

van den Dool, H. M., S. Saha, J. Schemm, and J. Huang (1997), A temporal interpolation method to obtain hourly atmospheric surface pressure tides in renalaysis 1979-1995, *J. Geophys. Res.*, *102*(D18), 22,013–22,024.

van Lipzig, N. P. M., E. V. van Meijgaard, and J. Oerlemans (2002), The spatial and temporal variability of the surface mass balance in Antarctica: Results from a regional atmospheric climate model, *Int. J. Climatol.*, *22*, 1197–1217.

Vaughan, D. G., J. L. Bamber, M. Giovinetto, J. Russell, and A. P. R. Cooper (1999), Reassessment of net surface mass balance in Antarctica, *J. Climate*, *12*, 933–946.

Velicogna, I., and J. Wahr (2002), A method for separating Antarctic postglacial revound and ice mass balance using future ICESat Geoscience Laser Altimeter System, Gravity Recovery and Climate Experiment, and GPS satellite data, *J. Geophys. Res.*, *107*(B10), 2263, doi:10.1029/2001JB000708.

Velicogna, I., J. Wahr, and H. M. van den Dool (2001), Can surface pressure be used to remove atmospheric contributions from GRACE data with sufficient accuracy to recover hydrological signals?, *J. Geophys. Res.*, *106*(B8), 16,415–16,434.

Vorob'ev, V. V., and T. G. Krasil'nikova (1994), Estimation of the accuracy of the atmospheric refractive index recovery from doppler shift measurements at frequencies used in the NAVSTAR system, *Phys. Atmos. Ocean*, *29*, 602–609.

Wahr, J., M. Molenaar, and F. Bryan (1998), Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, *J. Geophys. Res.*, *103*(B12), 30,205–30,229.

Wahr, J., D. Wingham, and C. Bentley (2000), A method of combining ICESat and GRACE satellite data to constrain Antarctic mass balance, *J. Geophys. Res.*, *105*(B7), 16,279–16,294.

Wahr, J., S. Swenson, V. Zlotnicki, and I. Velicogna (2004), Time-variable gravity from GRACE: First results, *Geophys. Res. Lett.*, *31*, L1101, doi:10.1029/2004GL019779.

Wahr, J. M., S. R. Jayne, and F. O. Bryan (2002), A method of inferring changes in deep ocean currents from satellite measurements of time-variable gravity, *J. Geophys. Res.*, *107*(C12), 3218, doi:10.1029/2001JC001274.

Ware, R., M. Exner, D. Feng, M. Gobunov, K. Hardy, B. Herman, Y. Kuo, T. Meehan, W. Melbourne, C. Rocken, W. Schreiner, S. Sokolovskiy, F. Soltheim, X. Zou, R. Anthes, S. Businger, and K. Trenberth (1996), GPS sounding of the atmosphere for Low Earth Orbit: Preliminary results, *Bull. Amer. Meteor. Soc.*, *77*, 19–40.

Wickert, J., C. Reigber, G. Beyerle, R. König, C. Marquardt, T. Schmidt, L. Grunwaldt, R. Galas, T. K. Meehan, W. G. Melbourne, and K. Hocke (2001), Atmosphere sounding by GPS radio occultation: First results from CHAMP, *Geophys. Res. Lett.*, *28*, 3263–3266, doi:10.1029/2001GL013117.

Wickert, J., G. Beyerle, G. Hajj, V. Schwieger, and C. Reigber (2002), GPS radio occultation with CHAMP: Atmospheric profiling utilizing the space-based single difference technique, *Geophys. Res. Lett.*, *29*(8), doi:10.1029/2001GL013982.

Wunsch, C., and D. Stammer (1995), The global frequency-wavenumber spectrum of oceanic estimated from TOPEX/POSEIDON altimetric measurements, *J. Geophys. Res.*, *100*(C12), 24,895–24,910.

Wunsch, C., and D. Stammer (1997), Atmospheric loading and the oceanic "inverted barometer" effect, *Rev. Geophys.*, *35*, 79–107.

Xie, J., S. Ge, S.-C. Han, and C. K. Shum (2004), Implementation of parallel least square algorithm for gravity field inversion, *Journal of Geospatial Engineering*, *6*(1), 47–57.

Yoder, C. F., J. G. Williams, J. O. Dickey, B. E. Schutz, R. J. Eanes, and B. D. Tapley (1983), Secular variation of Earth's gravitational harmonic J_2 coefficient from Lageos and nontial acceleration of Earth rotation, *Nature*, *303*, 757–762.

Zou, X., F. Vandenberghe, B. Wang, M. E. Gorbunov, Y.-H. Kuo, S. Sokolovskiy, J. C. Chang, J. G. Sela, and R. A. Anthes (1999), A ray-tracing operator and its adjoint for the use of GPS/MET refraction angle measurements, *J. Geophys. Res.*, *104*(D18), 22,301–22,318.

Zou, X., R. A. Anthes, H. Shao, J. C. Chang, and Y.-J. Zhu (2004), Impact of CHAMP radio occultation observations on global analysis and forecasts in the absence of AMSU radiance data, *J. Meteorol. Soc. Jpn.*, *82*, 533–549.